

PATENT ABSTRACTS OF JAPAN

(11)Publication number: 2002-135707

(43)Date of publication of application: 10.05.2002

---

(51)Int.Cl. H04N 5/91

G11B 27/031

H04N 7/173

---

(21)Application number: 2000-320450 (71)Applicant: BROTHER IND LTD

(22)Date of filing: 20.10.2000 (72)Inventor: KONISHI SHIROU

FUKAYA HIROSUKE

ISHIMOTO SEKI

TAGAWA NORIO

ICHIKAWA YUKIHISA

---

(54) VIDEO EDITING SYSTEM

(57)Abstract:

PROBLEM TO BE SOLVED: To a video editing system capable of shortening a time required from decision of performing editing work until start of the work.

SOLUTION: A server 20 prepares editing information such as thumbnail images and a clip starting point/ending point of video pictures in a video-tape 12 played back by a video cassette recorder 11, to store it in an editing information storage 32. A client 40 receives desired editing information from a server 20 to perform editing work based on the editing information. After finishing editing work, the client 40 instructs the editing

to the server 20. The server 20 reads out a video/audio stream from a video/audio stream storage 31 to perform the editing following the editing instruction and records in a DVD-R with a DVD writer 13. Since the editing information transmitted from the server 20 to the client 40 is smaller in data amount than the original image, the communication time is short, thereby the system enables to shorten the time required from decision of performing editing work until start of the work.

---

**LEGAL STATUS** [Date of request for examination] 17.01.2007

[Date of sending the examiner's decision of rejection]

[Kind of final disposal of application other than the examiner's decision of rejection or application converted registration]

[Date of final disposal for application]

[Patent number]

[Date of registration]

[Number of appeal against examiner's decision of rejection]

[Date of requesting appeal against examiner's decision of rejection]

[Date of extinction of right]

**\* NOTICES \***

**JPO and INPIT are not responsible for any  
damages caused by the use of this translation.**

1. This document has been translated by computer. So the translation may not reflect the original precisely.

2.\*\*\*\* shows the word which can not be translated.

3.In the drawings, any words are not translated.

---

## CLAIMS

---

[Claim(s)]

[Claim 1] It has the server, and this server and the client which can be communicated. Said server An edit information storage means to memorize the edit information for editing an image, and a Request-to-Send signal receiving means to receive the Request-to-Send signal transmitted from said client, When said Request-to-Send signal is received by this Request-to-Send signal receiving means It has an edit information transmitting means to transmit to said client by making edit information corresponding to the contents shown by said received Request-to-Send signal among the edit information memorized by said edit information storage means into partial edit information. A Request-to-Send signal transmitting means to transmit the Request-to-Send signal which shows the demand of which part said client receives as partial edit information among said edit information memorized by said edit information storage means to said server, The image edit system characterized by having an edit information receiving means to receive said partial edit information transmitted from said server, and an edit means to edit the edit information corresponding to a desired edit image based on the partial edit information received with this edit information receiving means.

[Claim 2] It has the server, and this server and the client which can be communicated. Said server The setting means which set up whether it would transmit to said client by making which part into partial edit information among the edit information memorized by an edit information storage means to memorize the edit information for editing an image, and this edit information storage means, A Request-to-Send signal receiving means to receive the Request-to-Send signal transmitted from said client, When said Request-to-Send signal is received by this Request-to-Send signal receiving means It has an edit information transmitting means to transmit to said client by making into partial edit information edit information corresponding to the contents set up by said setting means among the edit information memorized by said edit information storage means. A Request-to-Send signal transmitting means to transmit the Request-to-Send signal which shows the demand which receives the edit information corresponding to the contents to which said client was set by said setting means as

partial edit information to said server, The image edit system characterized by having an edit information receiving means to receive said partial edit information transmitted from said server, and an edit means to edit the edit information corresponding to a desired edit image based on the partial edit information received with this edit information receiving means.

[Claim 3] The inside of the edit information said server is further remembered to be by said edit information storage means, It has the setting means which set up whether it would transmit to said client by making which part into partial edit information. Said edit information transmitting means The image edit system according to claim 1 characterized by transmitting to said client by making into partial edit information edit information corresponding to the contents set up by said setting means among the edit information memorized by said edit information storage means.

[Claim 4] Said edit information transmitting means is ready-for-sending ability about two or more kinds of partial edit information corresponding to the contents of said received Request-to-Send signal. Said server It has the weighting means which carries out [ said ] weighting of two or more which classes are transmitted among classes corresponding to the receiving frequency for every contents of said received Request-to-Send signal. Furthermore, said setting means The image edit system according to claim 2 or 3 characterized by the ability to change the contents set up based on weighting by said weighting means.

[Claim 5] Said edit information transmitting means is the image edit system of any one publication of claim 1 which chooses it as order with small amount of information from said partial edit information, and is characterized by being ready-for-sending ability to said client thru/or claim 4.

[Claim 6] Said client is equipped with a change-request signal transmitting means to transmit the change-request signal which requires further modification of the contents set up by said setting means. Said server A change-request signal receiving means to receive the change-request signal transmitted from said client, The image edit system of any one publication of claim 2 characterized by having a setting Make Changes means to change the contents set up by said setting means based on the contents shown by the change-request signal which this change-request signal receiving means received thru/or claim 5.

[Claim 7] Said server is the image edit system of any one publication of claim 1 characterized by having further the edit information generation section which makes edit information from an image thru/or claim 6.

[Claim 8] Said edit information generation section is an image edit system according

to claim 7 characterized by generating the edit information corresponding to the contents set up by said setting means, and memorizing the generated edit information for said edit information storage means when said Request-to-Send signal receiving means receives said Request-to-Send signal transmitted from said client.

[Claim 9] An edit indication signal receiving means by which said server receives further the edit indication signal transmitted from said client. The image which corresponds from the image memorized by said image storage means according to the edit indication signal received by image storage means to memorize an image, and said edit indication signal receiving means is edited. It has an image transmitting means to transmit an edited image to said client. Said client Furthermore, the image edit system of any one publication of claim 1 characterized by having an edit indication signal transmitting means to transmit the edit indication signal corresponding to the edit result by said edit means to said server thru/or claim 8.

[Claim 10] Said image transmitting means is an image edit system according to claim 9 characterized by outputting said corresponding image to the image storage connected to said server.

---

#### DETAILED DESCRIPTION

---

[Detailed Description of the Invention]

[0001]

[Field of the Invention] This invention relates to the image edit system into which an

image is edited between a server and a client.

[0002]

[Description of the Prior Art] Drawing 16 thru/or drawing 18 are the explanatory views showing an example of the edit technique of the animation currently performed from the former. The technique of dividing one animation into some clips (Clip: scene) ( drawing 16 ), changing the playback sequence of those clips, or deleting ( drawing 17 ) and the unnecessary part of the created clip, and creating a new clip conventionally, is used ( drawing 18 ). Moreover, when using those technique, the technique of expressing and editing the die length of a clip on GUI (Graphical User Interface) is known.

[0003]

[Problem(s) to be Solved by the Invention] However, by the above-mentioned conventional technique, although required amount of information changes with extent of edit, all animation information must be prepared. For example, if there is no image of all the coma in the corresponding clip, an editing task is difficult [ when performing the easy editing task which replaces a certain clip and another clip, if the image of the head part of each clip is known, it is possible to perform an editing task but ] to delete the garbage in a certain clip. Moreover, when performing the above-mentioned editing task between the server connected through the network, and a client, since a client needs to receive all animation information from a server in advance of an editing task, a transfer of the file from a server to a client takes long duration to it. For this reason, even if a client side is the case where he wants to perform an easy editing task, it cannot start an editing task until a transfer of a file is completed (until it acquires all also including the information which does not have the need in edit). Moreover, when a client side wants to refer to the rear part of a clip, the thing which is the need is the rear part of a clip, but an editing task cannot be started until a transfer of all the files from a head to the rear is completed. Since the network band is narrow as compared with LAN in a company etc. and a transfer of a file takes long duration especially when performing an editing task via the telephone line, editing-task time amount will increase further and communication link cost will also increase.

[0004] Then, this invention is made in order to solve many above-mentioned problems, and it aims at realizing the image edit system which can shorten the time amount which it takes after determining to perform an editing task before actually starting an editing task.

[0005]

[The means for solving a technical problem, an operation, and an effect of the

[invention] In order that this invention may attain the above-mentioned purpose, in invention according to claim 1 It has the server, and this server and the client which can be communicated. Said server An edit information storage means to memorize the edit information for editing an image, and a Request-to-Send signal receiving means to receive the Request-to-Send signal transmitted from said client, When said Request-to-Send signal is received by this Request-to-Send signal receiving means It has an edit information transmitting means to transmit to said client by making edit information corresponding to the contents shown by said received Request-to-Send signal among the edit information memorized by said edit information storage means into partial edit information. A Request-to-Send signal transmitting means to transmit the Request-to-Send signal which shows the demand of which part said client receives as partial edit information among said edit information memorized by said edit information storage means to said server, The technical means of having had an edit information receiving means to receive said partial edit information transmitted from said server, and an edit means to edit the edit information corresponding to a desired edit image based on the partial edit information received with this edit information receiving means are used.

[0006] The edit information-storage means with which the server was equipped has memorized the edit information for editing an image, and if a client transmits to a server the Request-to-Send signal which shows the demand of which part to receive as partial edit information among the above-mentioned edit information, a server will transmit to a client by making the edit information corresponding to the contents shown by the received Request-to-Send signal into partial edit information. And a server edits the edit information corresponding to a desired edit image based on the received partial edit information. That is, since a client can receive edit information required in order to edit an image from a server and can be edited using the edit information, it needs to receive no images. Therefore, the time amount which it takes after determining to perform an editing task before actually starting an editing task can be shortened rather than the technique of performing an editing task like before, after a client receives all images. Moreover, since the communication link time amount between server clients can be shortened, communication link cost can also be reduced. Moreover, a server can transmit the edit information corresponding to the contents shown by the Request-to-Send signal transmitted from the client to a client. Therefore, a client can receive the edit information for which it wishes by changing a Request-to-Send signal, when it is not what the edit information received from the server wishes.

[0007] In invention according to claim 2, it has the server, and this server and the client which can be communicated. Said server The setting means which set up whether it would transmit to said client by making which part into partial edit information among the edit information memorized by an edit information storage means to memorize the edit information for editing an image, and this edit information storage means, A Request-to-Send signal receiving means to receive the Request-to-Send signal transmitted from said client, When said Request-to-Send signal is received by this Request-to-Send signal receiving means It has an edit information transmitting means to transmit to said client by making into partial edit information edit information corresponding to the contents set up by said setting means among the edit information memorized by said edit information storage means. A Request-to-Send signal transmitting means to transmit the Request-to-Send signal which shows the demand which receives the edit information corresponding to the contents to which said client was set by said setting means as partial edit information to said server, The technical means of having had an edit information receiving means to receive said partial edit information transmitted from said server, and an edit means to edit the edit information corresponding to a desired edit image based on the partial edit information received with this edit information receiving means are used.

[0008] That is, since a server can set up whether it transmits to a client by making which part into partial edit information among the edit information memorized by the edit information storage means, when the class and the contents of edit information which are demanded by each client differ from each other, the processing of generation of edit information, management, etc. in a server does not complicate it.

[0009] In invention according to claim 3, it sets to an image edit system according to claim 1. Said server Furthermore, it has the setting means which set up whether it would transmit to said client by making which part into partial edit information among the edit information memorized by said edit information storage means. The technical means of transmitting to said client by making into partial edit information edit information corresponding to the contents set up by said setting means among the edit information memorized by said edit information storage means are used for said edit information transmitting means.

[0010] That is, a server can also be transmitted to a client by making into partial edit information edit information corresponding to the contents which could transmit to the client by having made into partial edit information edit information corresponding to the contents shown by the Request-to-Send signal transmitted from the client, or

self set up. Therefore, a client can make the edit information for which it wishes transmit to a client by being able to save the time and effort which changes the class of Request-to-Send signal according to a situation when following the contents set as the server, and changing a Request-to-Send signal, when it is not what the edit information corresponding to the contents set as the server wishes.

[0011] In invention according to claim 4, it sets to an image edit system according to claim 2 or 3. Said edit information transmitting means Corresponding to the contents of said received Request-to-Send signal, it is ready-for-sending ability about two or more kinds of partial edit information. Said server It has the weighting means which carries out [ said ] weighting of two or more which classes are transmitted among classes corresponding to the receiving frequency for every contents of said received Request-to-Send signal. Furthermore, said setting means The technical means that the contents set up can be changed based on weighting by said weighting means are used.

[0012] Weighting of the server is carried out corresponding to the receiving frequency for every contents of the Request-to-Send signal which received, and it can change the contents set up based on the weighting. That is, since the edit information to which the frequency which a client uses in order to set up more highly than the edit information of other classes the priority of transmission of the edit information corresponding to the contents in a server when the frequency where the use frequency of the edit information of a specific class becomes high in a client, and the contents same as contents of the Request-to-Send signal are set up according to it becomes high became high is preferentially receivable, it is convenient.

[0013] In invention according to claim 5, in the image edit system of any one publication of claim 1 thru/or claim 4, said edit information transmitting means is chosen as order with small amount of information from said partial edit information, and the technical means that it is ready-for-sending ability are used for it to said client.

[0014] That is, since the contents of the image can be specified in many cases even if it does not necessarily transmit image data with much amount of information, such as data in which the start point and the ending point of a clip are shown, a server transmits partial edit information to order with small amount of information to a client. Therefore, possibility that communication link time amount is shortened and communication link cost can be reduced rather than the case where it transmits from the partial edit information that amount of information is large can be raised.

[0015] In the image edit system of any one publication of claim 2 thru/or claim 5 in

invention according to claim 6 said client It has a change-request signal transmitting means to transmit the change-request signal which requires modification of the contents set up by said setting means. Furthermore, said server A change-request signal receiving means to receive the change-request signal transmitted from said client, The technical means of having had a setting Make Changes means to change the contents set up by said setting means based on the contents shown by the change-request signal which this change-request signal receiving means received are used.

[0016] That is, a client can change the above-mentioned contents of a setting by transmitting a change-request signal to a server to change the contents of a setting in a server.

[0017] In invention according to claim 7, the technical means of having had further the edit information generation section which makes edit information from an image are used for said server in the image edit system of any one publication of claim 1 thru/or claim 6.

[0018] That is, in order to make edit information in a server, it is not necessary to make edit information in a client. Therefore, since a client does not need to hold the computer program for making edit information, it can realize the image edit system which it is flexible and can reduce cost required for edit.

[0019] In invention according to claim 8, in an image edit system according to claim 7, when said edit information generation section receives said Request-to-Send signal transmitted from said client with said Request-to-Send signal receiving means, it generates the edit information corresponding to the contents set up by said setting means, and the technical means of memorizing the generated edit information for said edit information storage means are used for it.

[0020] That is, since a server generates the edit information corresponding to the contents of a setting when it receives a Request-to-Send signal from a client, even if it is the case where the client is not demanding, as compared with the case where edit information is generated, it can utilize efficiently the storage region of the edit information in an edit information storage means, and can exclude the useless processing for generation and storage.

[0021] In the image edit system of any one publication of claim 1 thru/or claim 8 in invention according to claim 9 said server Furthermore, an edit indication signal receiving means to receive the edit indication signal transmitted from said client, The image which corresponds from the image memorized by said image storage means according to the edit indication signal received by image storage means to memorize

an image, and said edit indication signal receiving means is edited. It has an image transmitting means to transmit an edited image to said client, and the technical means of having had further an edit indication signal transmitting means to transmit the edit indication signal corresponding to the edit result by said edit means to said server are used for said client.

[0022] That is, a server memorizes an image, and since the image which corresponds according to the edit directions from a client is transmitted to a client, a client can perform image edit. Therefore, a client can create the medium by which the edit image was memorized by memorizing an edit image to the storage formed or connected to self.

[0023] In invention according to claim 10, the technical means of outputting said corresponding image to the image storage connected to said server are used for said image transmitting means in an image edit system according to claim 9.

[0024] That is, in order for a server not to transmit the corresponding image to a client but to transmit to the image store connected to self, a client does not need to memorize the image transmitted from the server. Therefore, a client does not need to secure the storage region for memorizing the received image while not requiring the communication link cost for receiving the image transmitted from the server. Moreover, a client can obtain the image which self edited by receiving from a server the storages (for example, DVD-R, DVD-RW, a video tape, etc.) with which the above-mentioned image store was equipped.

[0025]

[Embodiment of the Invention] Hereafter, the operation gestalt of the image edit system concerning this invention is explained with reference to drawing. In addition, with each operation gestalt shown below, the image recorded on the video tape is edited, and the case where the edit image is recorded on DVD-R is mentioned as an example, and is explained.

The main configuration and main functions of an image edit system concerning this operation gestalt are explained to the [main configuration and main functional] beginning with reference to drawing 1 and drawing 2. Drawing 1 is the explanatory view showing the main configurations of an image edit system with a block, and drawing 2 is the explanatory view showing with a block the main functions of the image edit system shown in drawing 1. In addition, each function shown with a block in drawing 2 shows the function realized when CPU (illustration abbreviation) performs a computer program. Moreover, each storage section functions with storages (illustration abbreviation), such as a hard disk and RAM.

[0026] As shown in drawing 1, the image edit system 10 consists of clients 40 connected with the server 20 and the server 20 via the network 14. The videocassette recorder 11 and the DVD (Digital Versatile Disc) writer 13 are connected to the server 20. The DVD writer 13 has the function which records an image and voice on DVD-R, and the function which reproduces DVD-R or DVD-ROM. A server 20 is equipped with a video capture card 21, the storage section 30, the stream edit activation section 22, and network I/F23. The storage section 30 is equipped with the edit information storage section 32 which memorizes the information (it is hereafter called edit information for short) for editing an image with an image and the voice stream storage section 31. Moreover, a network 14 is a network which combined a public switched network and the Internet.

[0027] A video capture card 21 incorporates the image and voice of a video tape 12 which were reproduced with the videocassette recorder 11, and by the capture card control unit 24 which is the software which operates a video capture card 21, the image and voice which were incorporated are changed into the stream data which can be processed by computer, and are memorized by an image and the voice stream storage section 31 (drawing 2). The edit information creation section 26 which is the software which creates edit information creates edit information using the image and voice stream memorized by an image and the voice stream storage section 31 (drawing 2), and memorizes the created edit information in the edit information storage section 32. Weighting of the edit information is carried out by predetermined conditions for every class, and it is hierarchized and memorized (drawing 1). The stream edit activation section 22 edits the image and voice stream memorized by an image and the voice stream storage section 31 according to the edit directions received from the client 40 through network I/F23, and records edited its image and voice stream on DVD-R with the DVD writer 13.

[0028] A client 40 is equipped with a personal computer (it is hereafter called PC for short) 42, a display 43, a keyboard 44, a mouse 45, and network I/F41 (drawing 1). A client 40 transmits the signal (it is hereafter called a Request-to-Send signal for short) which shows the Request to Send of edit information to a server 20 by the data transceiver section 50 (drawing 2) which is the software which transmits and receives data, and receives the edit information corresponding to the contents shown by the Request-to-Send signal from a server 20. This received edit information is decoded by the edit information decode section 48 (drawing 2) which decodes edit information and which is software, and that decode result is displayed on a display 43 by the edit-display display 46 which is the software which displays an edit display on a

display 43. Moreover, the above-mentioned decode result is memorized by the contents storage section 49 ( drawing 2 ) of edit.

[0029] On the edit display displayed on the display 43, the edit-display control unit 47 is the software for performing an editing task, and is well-known software. For example, the display gestalt of the edit display is a display gestalt of the Windows (Windows is trademark of Microsoft Corp.) method which is OS which Microsoft Corp. of the U.S. developed. According to this display gestalt, a screen top can be moved for a pointer with a mouse 45, and the easy actuation which carries out click actuation can perform an editing task. The edit result is memorized by the contents storage section 49 ( drawing 2 ) of edit, and the edit directions creation section 49 ( drawing 2 ) creates the signal (it is hereafter called an edit indication signal for short) for directing to a server 20 to perform edit corresponding to the contents of edit memorized by the contents storage section 49 of edit. And the stream edit activation section 22 ( drawing 1 ) of a server 20 reads an image and a voice stream from an image and the voice stream storage section 31 according to the edit indication signal transmitted from the client 40, and performs edit.

[0030] [The class of edit information], next the class of edit information are explained with reference to drawing 3 . Drawing 3 is the explanatory view showing the class of edit information memorized by the edit information storage section 32.

(1) Reduce the magnitude of the original image with a thumbnail image 32a thumbnail image. After a well-known algorithm's being able to generate a thumbnail image, for example, removing a high frequency component for a former image through a low-pass filter, an image is reducible by thinning out a pixel. If it is I-frame of MPEG or JPEG, a thumbnail can be easily created by extracting DC component. For this reason, although image quality (resolution) is inferior to a former image, since one coma of an image can be seen, the contents of the image can fully be grasped.

[0031] When performing a cut and rearrangement in a clip unit, in order to see and judge the frame the knot of a clip, i.e., the head of a clip, and near a tail as an editing task, it is desirable to transmit to a client from a server by making the thumbnail image of the frame into edit information. Moreover, the usual video images are 30 frames per second, and two continuous frames do not much have a difference, when the change (boundary of a clip) of a scene is removed. Therefore, when transmitting the thumbnail of the frame of the same number is considered, the direction which transmitted the frame to every spacing (for example, 1 second) of a certain is useful to grasping the whole clip from the continuous frame. Moreover, size of data can be made small. For example, a former image is what carried out 720x480JPEG

compression although the amount of data of a resolution 720x480RGB format (BMP) was 1,036,854 bytes, and is 67,289 bytes, and it becomes about 6.5% of the amount of data of a former image. Moreover, the resolution of a former image is changed, and in the JPEG compression set to 360x240, it is 21,478 bytes and becomes about 2.1% of the amount of data of a former image. Furthermore, in 180x120JPEG compression, it is 7,399 bytes and becomes about 0.7% of the amount of data of a former image.

[0032] (2) binary image 32b -- since the screen can fully be recognized even if it uses not a full color image but a monochrome binary image as a thumbnail, the contents of the image can fully be grasped. Although 24-bit data are needed to 1 dot if full color when it thinks by the simple bitmapped image, it ends with 1 bit by the binary image. Therefore, the amount of data of a thumbnail can be reduced to 1/24. Moreover, the amount of data can be further reduced by compressing binary image data. A binary image is generable by assigning "0", when the value is larger than a certain threshold and "1" and a value are small about each pixel. For example, in the case of a color picture, it changes into gray scale first. In order to change a RGB color picture into gray scale, the following transformations are made to act about each pixel.

[0033] Although a threshold is set up to  $Y=0.299R+0.587G+0.144B$ , next the above-mentioned gray scale, in order to generate a good (it is easy to be visibility) binary image, how to calculate the above-mentioned threshold is important. Since that to which the alphabetic character is written in the ink to a binary thing, for example, white paper, with a black image from the first will produce two peaks in a brightness histogram, it should just set a threshold as the midpoint. On the other hand, although it is difficult for the image whose image from the first is not binary to calculate the optimal threshold, simply, it makes a threshold the maximum of the target image, and the middle value of the minimum value, and makes them binary.

[0034] (3) some images -- since the screen can fully be recognized even if it uses as a thumbnail the data which trimmed some 32c images, the contents of the image can fully be grasped. That is, when the part by which the contents of the image are characterized exists in some images, the contents of the clip with which the image is contained if it can try to carry out a part can be known. Drawing 4 is the explanatory view showing signs that some images are trimmed. As shown in this drawing, the rectangle A surrounded by the horizontal n(for example, 80 pixels) x length m (for example, 60 pixels) is taken out, and new bitmapped image B is generated. Although the resolution of Image B drops to several [ of Image A / 1 (for example, 1/4)/] at this time, since the part by which the contents of the image A are characterized is contained in that image B, compared with Image A, there is no inferiority in it. And the

amount of data can be reduced to several [ 1 (for example, 1/4)/]. some images -- 32q is generable by extracting only the pixel applicable to the range of a rectangle A among former images by giving the information on the rectangle A starting with trimming (a, b, m, n), as shown in drawing 4 .

[0035] (4) Since the screen can be recognized even if it uses a full color image or not a binary image but a line drawing as a thumbnail 32d of line drawings, the contents of the image can fully be grasped. Since a line drawing takes out the profile of a binary image, it can reduce the amount of data sharply rather than a binary image. Moreover, when there are few parts which can generate a line drawing (it is sparse), a chain code can express an image by the very small amount of data. Moreover, a line drawing is generable with three steps of binary-izing of an image, detection of the verge, and thinning.

[0036] (5) The (closed caption CC) 32e closed caption is embedded at the 21st scanning line of a video signal. The scene in a clip can be specified by using the data in which the title of a clip is shown as the closed caption. There is a thing with the function which includes the part in drawing and an MPEG stream with the capture equipment of video in the existing MPEG encoder. Moreover, since the MPEG decoder usually used is equipped with the function which takes out the above-mentioned part, a closed caption can be taken out because this extracts data.

[0037] (6) By using the information which can be read in the appearance of 32f video tape of tape label information as tape label information, when there are two or more video tapes, a tape can be identified. For example, the following information is made into text data. a part for formats (MiniDV, Hi8, Video8 and VHS, VHS-C, etc.) (less than 8 bytes), manufacture name (less than 20 bytes), name-of-article (less than 20 bytes), and tape length [--] (less than 4 bytes) -- attributes (MP, ME, etc.) (less than 8 bytes), the image transcription approaches (SP, LP, a criterion, 3 times, etc.) (less than 8 bytes), and the contents of a label sheet (less than 256 bytes). Moreover, a line feed code (2 bytes) goes into each line. The data of one tape are a maximum of 338 bytes, and though the data size of 32f of tape label information has ten video tapes, it is 3380 bytes in total. Since the completion of transmitting is carried out in about 0.4 seconds, this can be said to be small [ the amount of data ] in the communication band of 64kbps(es). The operator by the side of a server reads in the appearance of a video tape, inputs the read information with a text editor etc., and 32f of tape label information creates text data.

[0038] (7) Although 32g \*\* voice stream of sound information was encoded to MPEG-Audio, divide and hold a part for time amount with a head. For example, when

dividing and holding a part for for about 3 seconds, it can divide among an audio frame, and a part for the time amount nearest to 3 seconds is divided and held. By hearing the sound for [ with a head ] time amount, the contents of the image can fully be grasped.

\*\* The contents of the clip can be guessed by some clips or the sound volume of the whole average being shown, or min and the maximum sound volume being shown. For example, since average sound volume differs from the maximum sound volume considerably between clips when the clip which photoed the car race is recorded after the clip of the contents which photoed the scene in the quiet place of a gap, the contents of the clip can be guessed by getting to know those information.

[0039] When the average sound volume for every clip is shown, if 2 bytes and 8 bit sampling show supposing 16 bit sampling shows sound volume, it will become the data size of 1 byte. The size mentioned here concerns with the die length of a clip and is fixed, and the amount of data is very small. since [ however, ] it becomes difficult to select the description only by equalizing sound volume simply so that the die length of a clip becomes long -- a clip -- only a part is taken out very much. As 32s of sound-volume information on each clip, although various information can be considered, the sound volume of all the samples in a clip is equalized, for example, and it considers as the central value of the sound volume of the clip. moreover, the sound volume of the sample for head n seconds (for example, 3 seconds) of a clip can be equalized (or maximum -- taking), and it can also consider as the central value of the sound volume of the clip. furthermore, a sample is divided into every [ n seconds (for example, 5 seconds) ] within a clip, and each is equalized about the divided part (or maximum -- taking), and you may make it express distribution of the sound volume within a clip In this case, the amount of data becomes the die length (second) of a clip / 5 times. Furthermore, the frequency-division information on the band of a head audio frame may be created.

[0040] \*\* Spectrum drawing 5 of a sound is a graph which shows the spectrum of nine bands of a sound. An axis of abscissa is a frequency and an axis of ordinate is the energy of a sound. According to this graph, the frequency characteristics (height of = sound) of a sound can be visualized, and a tone can be grasped visually. Supposing the magnitude of data expresses one sample with 4 bytes, with a sampling frequency of 32kHz PCM data the amount of 1 second will become  $4 \text{ bytes} * 32,000 = 128,000$  bytes, but since it will end with  $4 \text{ bytes} * 9 = 36$  bytes supposing it expresses one band of the spectrum of a sound with 4 bytes, the amount of data is sharply reducible. It can ask for the spectrum of a sound by the technique of common knowledge, such as the

Fourier transform and wavelet transform.

[0041] (8) Since the die length of each clip can find and the contents of the clip can specify based on the die length which found, respectively if the start point and the ending point of each clip are known when the contents of the video image which serves as a material when performing the editing tasks (for example, insertion, deletion, a copy, rearrangement, etc.) of the start point and 32h clip unit of ending points of a clip can be grasped by the time basis, the editing task of a clip unit can carry out. Drawing 6 is the explanatory view of DS showing the start point and the ending point of a clip. Since it expresses a time second with 1 byte at a time, respectively and a frame number is expressed with 1 byte about one start point or an ending point, it is a total of 4 bytes of amount of data. For example, even when the whole video image used as a material consists of 100 clips, the sum total of the amount of data in which a start point and an ending point are shown is 4 bytes \* 100+1 byte = 404 bytes, and the amount of data is very small.

[0042] Here, an example of technique which detects the start point and the ending point of a clip is explained with reference to drawing 7. Drawing 7 is a flow chart which shows the flow of the processing at the time of a server 20 detecting the start point and the ending point of the clip as edit information. This technique is characterized by inserting blank images (a deep-black screen, pure white screen, etc.) in a part making into the break of a clip beforehand by ten-frame continuation. If a server 20 acquires the frame image which constitutes the video image used as a material (step 12 (it abbreviates to S hereafter)) and a blank image is detected (S14:Yes), it will add "+1" to the counter value of the counter which counts a blank image (S16). Then, a server 20 will set up the time amount corresponding to the blank image of the 10th frame as a break of a clip, if detection of a blank image and addition of counted value are repeated from acquisition of a frame image (S12-S16) and counted value amounts to "10" until the counted value of a counter amounts to "10" (S18:Yes) (S20).

[0043] then, the case where a server 20 carries out zero reset of the counter of a blank image (S22), and the following frame image exists -- (S26:Yes) -- S12-S22 are repeated again. Moreover, a server 20 carries out zero reset of the counter of (S14:No) and a blank image, when the acquired frame image is not a blank image (S24), and it stands by until it acquires the following black image (S26, S12, S14). Moreover, the technique of detecting the break point of \*\* image transcription time as other technique of detecting the start point and the ending point of a clip (for example, JP,11-46336,A). \*\* The technique of dividing the video image of a material for every fixed time amount from a head, and setting those each up as a clip automatically. It is

effective technique to carry out the clip division of this for every elapsed time. \*\* The technique of detecting the part of the fade-in included in the video image of a material, or fade-out, and setting the detected part as the break of a clip. This is usually effective technique to carry out a clip division for every change of a scene in order to perform fade-in or fade-out at the change of a scene. In this case, the same algorithm as detection of the above-mentioned blank image is applicable. \*\* The technique of setting up clips one by one and going, while the person in charge by the side of a server looks at the video image of a material according to the hope of a client. This has the advantage that becoming the clip division which a client does not mean decreases, although time and effort is taken. Moreover, if this technique is used after carrying out a clip division roughly by other technique, a clip division can be performed efficiently.

[0044] (9) Since it corresponds to the number of clip start points, or the number of ending points, the number of clips and the number of number 32i clips are called for by performing the algorithm which detects above-mentioned clip start point and ending point. Moreover, a clip number can be attached when a clip start point is detected. It can find clipping [ to wish ] by specifying the number of clips, or a clip number roughly first, and narrowing down gradually after that.

[0045] (10) Since the die length of the clip can be guessed from the frame number in a frame number 32j clip, it is effective technique when the contents can be grasped with the die length of a clip. Moreover, elapsed time can also be known by transmitting in the recorded sequence. Moreover, the frame number which can be expressed by 2 bytes is  $2^{16}-1=65,535$  frame, and this is equivalent to the die length for about 36 minutes, supposing frame rates are 30fps(es) ( $65,535/30 =$  about 36). Therefore, the amount of data required [ if all clips are less than about 36 minutes ] in order that the whole may express 100 clips becomes  $2 \text{ bytes} * 100 = 200 \text{ bytes}$ , and can be managed with the very small amount of data. A frame number counts up and generates the number of the frames in each clip. Moreover, the time amount of a clip is found from a clip start point and an ending point, and it can generate by carrying out the multiplication of the frame rate to the time amount.

[0046] (11) When it is grasped how much point changing [ scene ] is in the video image of a scene change mark 32k material, an editing task can be performed by specifying scene change mark.

[0047] (12) Since the part with the high activity of a motion vector in screen has many motions, it can predict on edit that it is an important case. Moreover, although it is thought that the scene where activity is extremely low is a still picture, it is the part

which failed in video photography in many cases. Then, the technique of finding an editing point can be used by detecting the height of the activity of a screen. MPEG (Motion Picture Experts Group) applied with DVD, the digital camera, etc. etc. -- since it is compressing using the inter-frame motion vector and the activity of a screen can be expressed with a picture compression method using the motion vector information, the activity can be used for discernment of an image.

[0048] Drawing 8 (A) is the explanatory view of an arrow head showing a motion of a screen, and drawing 8 (B) is the explanatory view showing a part of coordinate of a motion vector. For example, as shown in drawing 8 (A), an arrow head is displayed on the part which should display a thumbnail in an edit display. Thereby, a screen can show to which direction it is moving how much. For example, a left-hand side arrow head shows that the screen is moving to the upper right, and it turns out that a screen reduces a rate in the depth direction and is moving to it with the arrow head of right-hand side shorter than the arrow head. Moreover, the activity of an image can also be sensuously expressed by expressing the height of activity with the gradation (gradual change of a shade) of from red until blue, and using it as a background of an arrow head.

[0049] GOP (Group of Pictures) whose amount of data is the coding unit of MPEG every -- since it can express by dozens of bytes, it is very small. Moreover, a motion vector is generated by taking out from the stream by which MPEG compression was carried out. The format of an MPEG stream is specified by 13818-2 of an ISO standard, and can take out a motion vector by processing equivalent to the existing MPEG decoder. A motion vector is used for PB macro book of PB picture in MPEG. A motion vector does not exist in I picture (and macro book of Type I). Moreover, although a backward motion vector also exists in B type macro book of B picture, the motion vector of only the forward direction is used here not using it. Thus, since the vector which moves to all the macro books of all frames does not necessarily exist, the processing about a motion vector is calculated about the data of a certain unit length, and equalizes and uses it. For example, a unit length makes it 15 frames (it sets to 30fps and is 0.5 seconds) using GOP of MPEG.

[0050] A motion vector is obtained as 1 set of coordinates for every macro book, as shown in drawing 8 (B). The technique of taking out the central value of the motion vector in the image of a unit length from these vector values is shown below.

\*\* Ask for the sum of a motion vector. Since the sum of each motion vector is normalized by the number of the macro books with which a motion vector exists, it calculates an average value and asks for the sum of each motion vector. It can ask for

the typical motion vector which shows by this in which direction the whole screen is moving.

\*\* Ask for the square sum of a motion vector. At the simple sum, although the parallel displacement of a screen is detectable, in rotation, by the simple sum, a result is set to "0" and it cannot detect a motion, for example. Thus, when a simple direction cannot be taken out, it asks for activity only in magnitude. For that purpose, the square sum of the die length of each vector is normalized and used by the block count.

\*\* Take out the maximum of a motion vector. When a part of image has a motion, there are not so many blocks with a meaningful motion vector. In such a case, a characteristic value cannot be taken out if it normalizes by the block count. Then, if the motion vector which has the greatest magnitude as a meaningful motion vector is adopted as central value, in the above-mentioned case, it can respond. Moreover, there is also technique of choosing 20% of the block count from the large thing of the magnitude of a vector as the same technique as this. Furthermore, it may use combining this approach and the technique of \*\* and \*\*, and you may ask for the sum or the square sum from what has large vector value instead of the vector of all blocks.

[0051] (13) Display it instead of an image, using as a representation color the color currently that in the screen which exists 32n of monochrome images used, and the color which equalized the whole. [ most ] For example, in case a thumbnail is displayed side by side, by displaying the part which is not transmitted in the mere color which has relation in a screen rather than being blank and displaying shows the ambient atmosphere of an image. Only the monochrome image of especially the change of a scene is [ but ] identifiable. Moreover, if it is monochrome, it can express by the browser of a client side only by describing the value of a color in HTML (Hyper Text Markup Language). Since it can be made to display on a browser only by describing in a RGB format like #12A844 in case it describes to HTML at this time, as compared with the case where color data are transmitted as it is, data size is very small. concrete -- < -- table width -- = -- " -- 20 -- " -- height -- = -- " -- 20 -- " -- bgcolor -- = -- " -- # -- FF -- 0000 -- " -- > -- < -- tr -- > -- < -- td -- > -- < -- / -- td -- > -- < -- / -- tr -- > -- < -- / -- table -- > -- a total -- 74 -- a byte -- a browser -- inside -- 20 -- x -- 20 -- a pixel -- size -- being red -- a rectangular head -- displaying -- it can make . For example, an image is analyzed of a RGB component, the mode or the average is calculated from distribution of RGB, and monochrome is determined.

[0052] The main contents of processing which the [main contents of processing] next a server 20, and a client 40 perform are explained.

(Weighting) First, in case a server 20 transmits edit information to a client 40, the technique of carrying out weighting of the ranking of whether for priority to be given to which edit information and to transmit is explained. Although various kinds of edit information is memorized by the edit information storage section 32 of a server 20 as shown in drawing 3, priority table 32p shown in drawing 9 determines the priority at the time of transmission. Priority table 32p is set up for every client, matches user ID (at the example shown in drawing 9, it is USER000001), priority, the demand frequency from a client, and the count of transmission to a client, and is constituted. In the example shown in drawing 9, the 1st place of priority is a thumbnail, the demand frequency is 38 and the count of transmission has become 87. Priority interchanges by study based on demand frequency and the count of transmission.

[0053] (Flow of the main processings which a server 20 and a client 40 perform) Next, with reference to drawing 11 and drawing 12, it explains that the main processings which a server 20 and a client 40 perform flow. Drawing 11 is a flow chart which shows the flow of the main processings which a server 20 performs, and drawing 12 is a flow chart which shows the flow of the main processings which a client 40 performs. A server 20 generates various kinds of edit information mentioned above based on the image of the video tape 12 played with the videocassette recorder 11 (drawing 1), and memorizes it in the edit information storage section 32 (drawing 2) (S10 of drawing 11). Moreover, if a client 40 transmits the signal (a Request-to-Send signal is called hereafter) which shows the Request to Send of edit information to a server 20 (S500 of drawing 12), a server 20 will receive the Request-to-Send signal (S50 of drawing 11), will read the edit information (partial edit information is called hereafter) corresponding to the contents shown by the Request-to-Send signal from the edit information storage section 32, and will transmit to a client 40 (S100).

[0054] Here, with reference to the flow chart of drawing 14 which shows it, it explains that the partial edit information transmitting processing which a server 20 performs in S100 of drawing 11 flows. The Request-to-Send flag which shows that the server 20 received the Request-to-Send signal from the client 40 is checked (S102), and when there is a flag, read-out and its read partial edit information are transmitted for the partial edit information corresponding to the contents shown by (S104:Yes) and the Request-to-Send signal to a client 40 from the edit information storage section 32 (S106). Then, although a server 20 resets a Request-to-Send flag (S108) and a Request-to-Send flag is checked (S102), since it is reset in S108, it judges with having no Request-to-Send flag (S104>No). Then, if it judges with a server 20 not transmitting partial edit information, i.e., the communication line being vacant as for it

(S110:No), partial edit information will be transmitted (S112).

[0055] In addition, since, as for a thumbnail, usefulness changes with parts as mentioned above when the partial edit information to transmit is a thumbnail, in one clip, they are the head of \*\* clip, and the thumbnail of a tail. \*\* The thumbnail of a certain time amount (for example, 1 second) unit (this also transmits toward a center from a head and a tail). \*\* others -- a thumbnail. Moreover, if the transmitting sequence of a clip is automatic, it transmits from a head, but when there are directions of a client, a client makes high priority of a clip demanded at the end. And a server 20 checks a Request-to-Send flag, without waiting for termination of partial edit information (S102). That is, a Request to Send is promptly answered from a client, and the partial edit information corresponding to the Request to Send is transmitted. When the Request to Send from a client is detected and there is information under transmission temporarily according to the approach by said ranking, the transmission is canceled and priority is given to the directions from a client.

[0056] And a client 40 receives the partial edit information transmitted from the server 20 (S550 of drawing 12 ), displays the received partial edit information on a display 43, and performs an editing task using a keyboard 44 and a mouse 45 (S600). And a client 40 transmits an edit indication signal to a server 20, after ending an editing task (S650). Then, the edit indication signal transmitted from the client 40 is received (S150 of drawing 11 ), the stream editorial department 22 ( drawing 2 ) performs read-out and its edit for an image and a voice stream from an image and the voice stream storage section 31 based on the received edit indication signal, and a server 20 outputs the edit result to the DVD writer 13 (S200). Thereby, the DVD writer 13 writes the inputted edit result in DVD-R, and DVD-R on which the contents into which the client 40 edited the video tape 12 were recorded completes it.

[0057] Moreover, a server 20 changes the contents of a setting of priority corresponding to the frequency of the Request to Send from a client 40 (S250 of drawing 11 ). Here, with reference to the flow chart of drawing 13 which shows it, it explains that the setting Make Changes processing which a server 20 performs flows. If the Request-to-Send signal transmitted from the client 40 is received (S252:Yes), a server 20 will add a predetermined value to the demand frequency ( drawing 9 ) matched with the edit information specified by the received Request-to-Send signal (S254), and will update User Information (S256).

[0058] Then, a server 20 measures the demand frequency after adding a predetermined value, and the demand frequency of other edit information that priority is higher than self (S258). When other edit information on demand frequency lower

than the demand frequency of the edit information by which the demand was carried out [ above-mentioned ] exists, the priority of (S260:Yes), and the edit information besides the above of which the demand was done [ above-mentioned ] is replaced (S262). [ edit information, and ] That is, the demand frequency of a client can be made to reflect in priority by study. In addition, since it exists from the edit information which exists one per video tape to the edit information which exists one per clip, or the edit information whose one exists for every frame, the above-mentioned predetermined value changes with edit information. For example, about a monochrome image, a binary image, and a line drawing, a predetermined value is "1", respectively, and sound volume and a motion vector are "10", respectively, and are set to "50", respectively about a clip start point, an ending point, and the number of clips and a number.

[0059] Moreover, when the edit information of the same kind on other clips is required from a client while transmitting the edit information on a certain clip (for example, thumbnail), at the time of the next transmission, transmission of a clip including not edit information but the newly specified edit information transmitted before is started. Furthermore, since it is thought that the information on the clip is unnecessary any longer when the edit information is omitted by the editing task of a client while transmitting the edit information on a certain clip (for example, thumbnail), transmission can be suspended and other edit information can be transmitted. Moreover, since possibility that the information on the clip is needed is small when it is decided that the clip is adopted as an edit result, transmission of edit information can be suspended at the time. In addition, since the thumbnail to all frames is equal to transmitting a video image, it can be removed from priority.

[0060] [Effectiveness of an operation gestalt]

(1) As mentioned above, if the image edit system 10 of the above-mentioned operation gestalt is used, since a client 40 can receive edit information required in order to edit an image from a server 20 and can be edited using the edit information, it needs to receive no images. Therefore, the time amount which it takes after determining to perform an editing task before actually starting an editing task can be shortened rather than the technique of performing an editing task like before, after a client 40 receives all images. Moreover, since the communication link time amount between server clients can be shortened, communication link cost can also be reduced.

(2) Moreover, a server 20 can transmit the edit information corresponding to the contents shown by the Request-to-Send signal transmitted from the client 40 to a

client 40. Therefore, a client 40 can receive the edit information for which it wishes by changing a Request-to-Send signal, when it is not what the edit information received from the server 20 wishes. Moreover, a client 40 can save the time and effort which changes the class of Request-to-Send signal according to a situation, when following the contents set as the server 20.

[0061] (3) Further, a server 20 carries out weighting of the priority of transmission of edit information corresponding to Request-to-Send frequency, and can change the priority set up based on the weighting. That is, if the operating frequency of the edit information on a specific class becomes high in a client 40, in order to set up the priority of transmission of the edit information more highly than the edit information on other classes, since a client 40 can receive preferentially the edit information to which the frequency to be used became high, it is convenient [ a server 20 ].

(4) Furthermore, in order to make edit information in a server 20, it is not necessary to make edit information in a client 40. Therefore, since a client 40 does not need to hold the computer program for making edit information, it can realize the image edit system which can raise versatility and can reduce cost required for edit.

[0062] (5) And further, a server 20 memorizes an image and a voice stream, and in order to transmit the edit information which corresponds according to the edit directions from a client 40 to a client 40, a server 20 does not need to memorize an image. Therefore, a server 20 can mitigate the burden for memorizing an image and a voice stream.

(6) Further, a server 20 does not transmit the edited image to a client 40, and since it outputs to the DVD writer 13 connected to self, a client 40 does not need to memorize the image transmitted from the server 20. Therefore, a client 40 does not need to secure the storage region for memorizing the received image while not requiring the communication link cost for receiving the image transmitted from the server 20. Moreover, a client 40 can obtain the image which self edited by receiving DVD-R from a server 20.

[0063] Operation gestalt] besides [

(1) it mentioned above as edit information by which combination priority attachment of the edit information is carried out -- it is not independent, and as shown below, it can also combine and use. In this case, although usefulness becomes high especially with combination, it is effective if the amount of data uses the combination which does not increase so much.

\*\* According to this combination, grasp of the contents of the clip becomes easy the fade-in point in one clip and a fade-out point, and two thumbnails corresponding to

them.

\*\* Only by the thumbnail of a clip head, and the data of a motion vector motion vector, although usefulness is low, since an image can be guessed from a thumbnail and a motion, usefulness can be raised by adding the thumbnail of a clip head to the data of a motion vector.

[0064] \*\* Rather than sound information and a top thumbnail image, since there is little amount of data, a sound can make an image only one sheet and it can decrease the amount of data because others transmit only sound information.

\*\* a monochrome image -- and -- \*\*\*\*\* -- the contents of the clip can be sensuously made intelligible by coloring a line drawing representing the screen. The same effectiveness is acquired even if it displays a binary image not in black and white but in a representation color.

\*\* Since time and effort is taken if the field transmitted each time is specified when [ of a thumbnail and an image ] transmitting some images in part, it is natural to fix uniformly the field made into a thumbnail. In addition, when combining edit information as mentioned above, it is effective to change priority attachment of edit information which should transmit to a degree by what kind of edit information is transmitted before. For example, since usefulness is very high, as for the thumbnail of the head of a clip, it is desirable to control to transmit the thumbnail of the head of a clip first and to transmit a motion vector, a monochrome image, sound information, etc. after that.

[0065] (2) In a client 40, the above-mentioned priority can also be changed directly.

Drawing 10 is the explanatory view showing the contents of a display of a display 43 (drawing 1) prepared in the client 40. As shown in drawing 10, the study situation of the priority of edit information is displayed on a display 43. In the example shown in drawing 10, priority is displayed by the tabular format, the 1st place of priority is a frame number, and the 2nd place has become a thumbnail. Moreover, the check box which chooses whether it carries out clear [ of the above-mentioned study ] or it saves, and the carbon button (go!!) which decides selection are displayed on the lower part of a table. A change of priority can be made by the technique of dragging and moving the alphabetic character box which inputs an alphabetic character and which changes the figure which shows priority into an alphabetic character box. Moreover, if KURIYA is checked by actuation of a mouse 45 and go!! is pushed, the contents of the study situation currently displayed will be eliminated. Moreover, when preservation is checked and go!! is pushed, the contents of the study situation currently displayed are saved. In addition, in a client 40, when modification of priority or actuation of elimination is performed, the change-request signal corresponding to the actuation is

transmitted to a server 20. That is, a client 40 can change the above-mentioned contents of a setting by transmitting a change-request signal to a server 20 to change the contents of a setting in a server 20. In addition, the image edit system at the time of adopting this configuration is equivalent to the image edit system concerning claim 6.

[0066] (3) By the size of the amount of data of edit information, priority attachment of transmission of edit information can be performed, and it can also memorize in the edit information storage section 32. Drawing 15 is an explanatory view at the time of setting up priority for the small edit information on the amount of data highly. That is, since the contents of the image may be able to be specified even if it does not necessarily transmit image information with much amount of information, such as data in which a frame number, and the start point and the ending point of a clip are shown, etc., a server 20 transmits edit information to order with small amount of information to a client 40. Therefore, possibility that communication link time amount is shortened and communication link cost can be reduced rather than the case where it transmits from the large edit information on amount of information can be raised. In addition, the image edit system at the time of adopting this configuration is equivalent to the image edit system concerning claim 5.

[0067] (4) When the edit information generation section 26 of a server 20 receives the Request-to-Send signal transmitted from the client 40, it can generate the edit information corresponding to the contents of a setting, and it can also constitute it so that the generated edit information may be memorized in the edit information storage section 32. According to this configuration, since what is necessary is just to generate the edit information corresponding to the contents of a setting when a Request-to-Send signal is received from a client 40, even if it is the case where the client 40 is not demanding, as compared with the case where edit information is generated, a server 20 can utilize efficiently the storage region of the edit information in the edit information storage section 32, and can exclude the useless processing for generation and storage. In addition, the image edit system at the time of adopting this configuration is equivalent to the image edit system concerning claim 8.

[0068] (5) An edit image and voice can be transmitted to a client 40 from a server 20, and although the case where an edit image and voice were outputted to the DVD writer 13 connected to the server 20 from a server 20 was explained, it can also constitute from an above-mentioned operation gestalt so that it may write in image storage, such as a DVD writer connected to the storage or the client 40 with which the client 40 was equipped. According to this configuration, the record medium with

which an edit image and voice were recorded can be created in a client 40. In addition, the image edit system at the time of adopting this configuration is equivalent to the image edit system concerning claim 9.

(6) Although the case where one set of one set of a client 40 and a server 20 was connected via the network 14 as an image edit system concerning this invention was mentioned as the example and the above-mentioned operation gestalt explained it, of course, this invention is applicable also to the combination of one set of two or more clients and a server, one set of a client and two or more servers, or two or more clients and two or more servers.

[0069] The edit information storage section 32 of the [correspondence relation of each claim and operation gestalt] server 20 corresponds to the edit information storage means concerning claim 1, and the data transceiver section 27 corresponds to a Request-to-Send signal receiving means and an edit information transmitting means. Moreover, the data transceiver section 50 of a client 40 corresponds to the Request-to-Send signal transmitting means and edit information receiving means concerning claim 1, and the edit-display display 46, the edit-display control unit 47, the edit information decode section 48, and the contents storage section 49 of edit correspond to an edit means. Moreover, priority table 32p ( drawing 9 ) corresponds to the setting means concerning claim 2, and the edit information generation section 26 of a server 20 corresponds to the edit information generation section concerning claim 7. Furthermore, the data transceiver section 27 of a server 20 corresponds to the edit indication signal receiving means concerning claim 9, and the data transceiver section 50 of a client 40 corresponds to an edit indication signal transmitting means. Furthermore, the image and the voice stream storage section 31 of a server 20 correspond to the image storage means concerning claim 9, and the DVD writer 13 corresponds to the image storage concerning claim 10.

[0070] And S50 ( drawing 11 ) which a server 20 performs functions as a Request-to-Send signal receiving means concerning claim 1, and S100 functions as an edit information transmitting means. Moreover, S500 ( drawing 12 ) which a client performs functions as a Request-to-Send signal transmitting means concerning claim 1, S550 functions as an edit information receiving means, and S600 functions as an edit means. Furthermore, S252-S262 which a server 20 performs ( drawing 13 ) function as a weighting means concerning claim 4, it functions as an edit indication signal receiving means which requires S150 ( drawing 11 ) for claim 9, and S200 functions as an image transmitting means. Furthermore, S650 ( drawing 12 ) which a client 40 performs functions as an edit indication signal transmitting means.

---

## DESCRIPTION OF DRAWINGS

---

[Brief Description of the Drawings]

[Drawing 1] It is the explanatory view showing the main configurations of the image edit system concerning the operation gestalt of this invention with a block.

[Drawing 2] It is the explanatory view showing with a block the main functions of the image edit system shown in drawing 1.

[Drawing 3] It is the explanatory view showing the class of edit information memorized by the edit information storage section 32.

[Drawing 4] It is the explanatory view showing signs that some images are trimmed.

[Drawing 5] It is the graph which shows the spectrum of nine bands of a sound.

[Drawing 6] It is the explanatory view of DS showing the start point and the ending point of a clip.

[Drawing 7] It is the flow chart which shows the flow of the processing at the time of a server 20 detecting the start point and the ending point of the clip as edit information.

[Drawing 8] Drawing 8 (A) is the explanatory view of an arrow head showing a motion of a screen, and drawing 8 (B) is the explanatory view showing a part of coordinate of a motion vector.

[Drawing 9] It is the explanatory view showing the configuration of priority table 32p.

[Drawing 10] It is the explanatory view showing the contents of a display of a display 43 prepared in the client 40.

[Drawing 11] It is the flow chart which shows the flow of the main processings which a server 20 performs.

[Drawing 12] It is the flow chart which shows the flow of the main processings which a client 40 performs.

[Drawing 13] It is the flow chart which shows the flow of the setting Make Changes processing which a server 20 performs.

[Drawing 14] It is the flow chart which shows the flow of the edit information transmitting processing which a server 20 performs.

[Drawing 15] It is an explanatory view at the time of setting up priority for the small edit information on the amount of data highly.

[Drawing 16] It is the explanatory view showing an example of the edit technique of the animation currently performed from the former.

[Drawing 17] It is the explanatory view showing an example of the edit technique of the animation currently performed from the former.

[Drawing 18] It is the explanatory view showing an example of the edit technique of the animation currently performed from the former.

[Description of Notations]

10 Image Edit System

11 Videocassette Recorder

13 DVD Writer (Image Storage)

20 Server

31 Image and Voice Stream Storage Section (Image Storage Means)

32 Edit Information Storage Section (Edit Information Storage Means)

32p Priority table (setting means)

40 Client

(19)日本国特許庁 (JP)

## (12) 公開特許公報 (A)

(11)特許出願公開番号

特開2002-135707

(P2002-135707A)

(43)公開日 平成14年5月10日 (2002.5.10)

(51)Int.Cl.<sup>7</sup>  
H 04 N 5/91  
G 11 B 27/031  
H 04 N 7/173

識別記号  
6 1 0

F I  
H 04 N 7/173  
5/91  
G 11 B 27/02

テマコード(参考)  
6 1 0 A 5 C 0 5 3  
N 5 C 0 6 4  
B 5 D 1 1 0

審査請求 未請求 請求項の数10 O L (全 17 頁)

(21)出願番号 特願2000-320450(P2000-320450)

(22)出願日 平成12年10月20日 (2000.10.20)

(71)出願人 000005267  
プラザー工業株式会社  
愛知県名古屋市瑞穂区苗代町15番1号

(72)発明者 小西 伺郎  
名古屋市瑞穂区苗代町15番1号 プラザー  
工業株式会社内

(72)発明者 深谷 浩祐  
名古屋市瑞穂区苗代町15番1号 プラザー  
工業株式会社内

(74)代理人 100095795  
弁理士 田下 明人 (外1名)

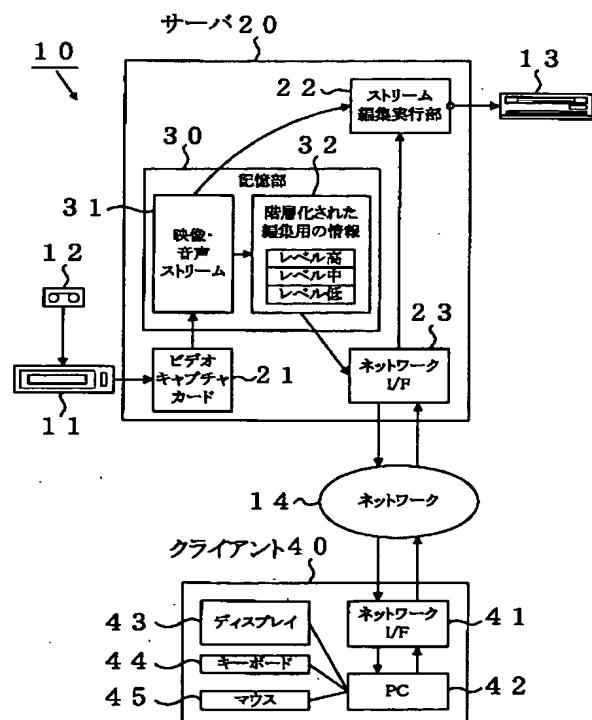
最終頁に統ぐ

### (54)【発明の名称】 映像編集システム

#### (57)【要約】

【課題】 編集作業を行うことを決定してから、実際に編集作業を開始するまでにかかる時間を短縮することができる映像編集システムを実現する。

【解決手段】 サーバ20は、ビデオデッキ11により再生されたビデオテープ12の映像のサムネイル画像、クリップ開始点・終了点などの編集情報を生成し、編集情報記憶部32に記憶する。クライアント40は、希望する編集情報をサーバ20から受信し、その編集情報に基づいて編集作業を行う。クライアント40は、編集作業が終了すると、編集指示をサーバ20に行い、サーバ20は、編集指示に従って映像・音声ストリーム記憶部31から映像・音声ストリームを読み出して編集を行い、DVDライタ13によりDVD-Rに記録する。サーバ20からクライアント40へ送信される編集情報は、元の映像よりもデータ量が小さいため、通信時間が短く、編集決定から開始までの時間を短縮できる。



## 【特許請求の範囲】

【請求項1】 サーバと、このサーバと通信可能なクライアントとを備えており、前記サーバは、映像を編集するための編集情報を記憶する編集情報記憶手段と、前記クライアントから送信された送信要求信号を受信する送信要求信号受信手段と、

この送信要求信号受信手段により前記送信要求信号が受信された場合に、前記編集情報記憶手段に記憶されている編集情報のうち、前記受信された送信要求信号により示される内容に対応した編集情報を部分編集情報として前記クライアントへ送信する編集情報送信手段とを備えており、

前記クライアントは、前記編集情報記憶手段に記憶されている前記編集情報のうち、どの部分を部分編集情報として受信するかの要求を示す送信要求信号を前記サーバへ送信する送信要求信号送信手段と、

前記サーバから送信されて来た前記部分編集情報を受信する編集情報受信手段と、この編集情報受信手段により受信した部分編集情報に基づいて所望の編集映像に対応する編集情報を編集する編集手段とを備えたことを特徴とする映像編集システム。

【請求項2】 サーバと、このサーバと通信可能なクライアントとを備えており、前記サーバは、

映像を編集するための編集情報を記憶する編集情報記憶手段と、

この編集情報記憶手段に記憶されている編集情報のうち、どの部分を部分編集情報として前記クライアントへ送信するかを設定した設定手段と、

前記クライアントから送信された送信要求信号を受信する送信要求信号受信手段と、

この送信要求信号受信手段により前記送信要求信号が受信された場合に、前記編集情報記憶手段に記憶されている編集情報のうち、前記設定手段により設定された内容に対応した編集情報を部分編集情報として前記クライアントへ送信する編集情報送信手段とを備えており、

前記クライアントは、前記設定手段により設定された内容に対応した編集情報を部分編集情報として受信する要求を示す送信要求信号を前記サーバへ送信する送信要求信号送信手段と、

前記サーバから送信されて来た前記部分編集情報を受信する編集情報受信手段と、

この編集情報受信手段により受信した部分編集情報に基づいて所望の編集映像に対応する編集情報を編集する編集手段とを備えたことを特徴とする映像編集システム。

【請求項3】 前記サーバは、さらに、前記編集情報記憶手段に記憶されている編集情報のうち、どの部分を部

分編集情報として前記クライアントへ送信するかを設定した設定手段を有し、

前記編集情報送信手段は、前記編集情報記憶手段に記憶されている編集情報のうち、前記設定手段により設定された内容に対応した編集情報を部分編集情報として前記クライアントへ送信することを特徴とする請求項1に記載の映像編集システム。

【請求項4】 前記編集情報送信手段は、前記受信された送信要求信号の内容に対応して複数種類の部分編集情報を送信可能であり、

前記サーバは、さらに、前記複数種類のうち、どの種類を送信するかを前記受信された送信要求信号の内容毎の受信頻度に対応して重み付けする重み付け手段を備えており、前記設定手段は、前記重み付け手段による重み付けに基づいて、設定されている内容を変更可能であることを特徴とする請求項2または請求項3に記載の映像編集システム。

【請求項5】 前記編集情報送信手段は、前記部分編集情報の中から、情報量の小さい順に選択して前記クライアントへ送信可能であることを特徴とする請求項1ないし請求項4のいずれか1つに記載の映像編集システム。

【請求項6】 前記クライアントは、さらに、前記設定手段により設定されている内容の変更を要求する変更要求信号を送信する変更要求信号送信手段を備えており、

前記サーバは、前記クライアントから送信された変更要求信号を受信する変更要求信号受信手段と、

この変更要求信号受信手段が受信した変更要求信号により示される内容に基づいて前記設定手段により設定されている内容を変更する設定内容変更手段とを備えたことを特徴とする請求項2ないし請求項5のいずれか1つに記載の映像編集システム。

【請求項7】 前記サーバは、さらに、映像から編集情報を作り出す編集情報生成部を備えたことを特徴とする請求項1ないし請求項6のいずれか1つに記載の映像編集システム。

【請求項8】 前記編集情報生成部は、前記クライアントから送信された前記送信要求信号を前記送信要求信号受信手段により受信した際に、前記設定手段により設定された内容に対応する編集情報を生成し、その生成した編集情報を前記編集情報記憶手段に記憶することを特徴とする請求項7に記載の映像編集システム。

【請求項9】 前記サーバは、さらに、前記クライアントから送信された編集指示信号を受信する編集指示信号受信手段と、

映像を記憶する映像記憶手段と、

前記編集指示信号受信手段により受信された編集指示信

号にしたがって前記映像記憶手段に記憶されている映像から該当する映像を編集し、編集済映像を前記クライアントへ送信する映像送信手段とを備えており、前記クライアントは、さらに、前記編集手段による編集結果に対応した編集指示信号を前記サーバへ送信する編集指示信号送信手段を備えたことを特徴とする請求項1ないし請求項8のいずれか1つに記載の映像編集システム。

【請求項10】前記映像送信手段は、前記該当する映像を前記サーバに接続された映像記憶装置へ出力することを特徴とする請求項9に記載の映像編集システム。

【発明の詳細な説明】

【0001】

【発明の属する技術分野】この発明は、サーバおよびクライアント間で映像の編集を行う映像編集システムに関する。

【0002】

【従来の技術】図16ないし図18は、従来から行われている動画の編集手法の一例を示す説明図である。従来は、1つの動画をいくつかのクリップ(Clip:場面)に分け(図16)、それらのクリップの再生順序を変更したり(図17)、作成したクリップの不要な部分を削除して新たなクリップを作成したりするなどの手法が用いられている(図18)。また、それらの手法を用いる場合、クリップの長さをGUI(Graphical User Interface)上に表して編集する手法が知られている。

【0003】

【発明が解決しようとする課題】しかし、上記従来の手法では、編集の程度により必要な情報量が異なるにもかかわらず、総での動画情報を用意しなければならない。たとえば、あるクリップと別のクリップとを入れ替えるだけの簡単な編集作業を行う場合は、それぞれのクリップの先頭部分の画像が分かれば編集作業を行うことが可能であるが、あるクリップ内の不要部分を削除したい場合には、該当するクリップ中の総てのコマの画像がなければ編集作業が困難である。また、上記編集作業をネットワークを介して接続されたサーバおよびクライアント間で行う場合は、クライアントは、編集作業に先だって総ての動画情報をサーバから受信する必要があるため、サーバからクライアントへのファイルの転送に長時間を要する。このため、クライアント側は、簡単な編集作業を行いたい場合であっても、ファイルの転送が完了するまで(編集に必要な情報も含めて総て取得するまで)編集作業を開始することができない。また、クライアント側は、クリップの後尾部分を参照したい場合において必要なのはクリップの後尾部分であるが、先頭から後尾までの総てのファイルの転送が終了するまで編集作業を開始することができない。特に、電話回線経由で編集作業を行う場合は、企業内のLANなどと比較してネットワークの帯域が狭いため、ファイルの転送に長時間

を要するので、編集作業時間はより一層増加するし、通信コストも増大してしまう。

【0004】そこで、この発明は、上記諸問題を解決するためになされたものであり、編集作業を行うことを決定してから、実際に編集作業を開始するまでにかかる時間を短縮することができる映像編集システムを実現することを目的とする。

【0005】

【課題を解決するための手段、作用および発明の効果】

10 この発明は、上記目的を達成するため、請求項1に記載の発明では、サーバと、このサーバと通信可能なクライアントとを備えており、前記サーバは、映像を編集するための編集情報を記憶する編集情報記憶手段と、前記クライアントから送信された送信要求信号を受信する送信要求信号受信手段と、この送信要求信号受信手段により前記送信要求信号が受信された場合に、前記編集情報記憶手段に記憶されている編集情報のうち、前記受信された送信要求信号により示される内容に対応した編集情報を部分編集情報として前記クライアントへ送信する編集情報送信手段とを備えており、前記クライアントは、前記編集情報記憶手段に記憶されている前記編集情報のうち、どの部分を部分編集情報として受信するかの要求を示す送信要求信号を前記サーバへ送信する送信要求信号送信手段と、前記サーバから送信されて来た前記部分編集情報を受信する編集情報受信手段と、この編集情報受信手段により受信した部分編集情報に基づいて所望の編集映像に対応する編集情報を編集する編集手段とを備えたという技術的手段を用いる。

【0006】サーバに備えられた編集情報記憶手段は、映像を編集するための編集情報を記憶しており、クライアントが、上記編集情報のうち、どの部分を部分編集情報として受信するかの要求を示す送信要求信号をサーバへ送信すると、サーバは、受信した送信要求信号により示される内容に対応した編集情報を部分編集情報としてクライアントへ送信する。そして、サーバは、受信した部分編集情報に基づいて所望の編集映像に対応する編集情報を編集する。つまり、クライアントは、映像を編集するために必要な編集情報をサーバから受信し、その編集情報により編集を行うことができるため、総ての映像を受信する必要がない。したがって、従来のように、クライアントが総ての映像を受信してから編集作業を行う手法よりも、編集作業を行うことを決定してから実際に編集作業を開始するまでにかかる時間を短縮することができる。また、サーバ・クライアント間の通信時間を短縮できるため、通信コストを低減することもできる。また、サーバは、クライアントから送信された送信要求信号により示される内容に対応した編集情報をクライアントへ送信することができる。したがって、クライアントは、サーバから受信した編集情報が希望するものでない場合は、送信要求信号を変更することにより、希望する

編集情報を受信することができる。

【0007】請求項2に記載の発明では、サーバと、このサーバと通信可能なクライアントとを備えており、前記サーバは、映像を編集するための編集情報を記憶する編集情報記憶手段と、この編集情報記憶手段に記憶されている編集情報のうち、どの部分を部分編集情報として前記クライアントへ送信するかを設定した設定手段と、前記クライアントから送信された送信要求信号を受信する送信要求信号受信手段と、この送信要求信号受信手段により前記送信要求信号が受信された場合に、前記編集情報記憶手段に記憶されている編集情報のうち、前記設定手段により設定された内容に対応した編集情報を部分編集情報として前記クライアントへ送信する編集情報送信手段とを備えており、前記クライアントは、前記設定手段により設定された内容に対応した編集情報を部分編集情報として受信する要求を示す送信要求信号を前記サーバへ送信する送信要求信号送信手段と、前記サーバから送信されて来た前記部分編集情報を受信する編集情報受信手段と、この編集情報受信手段により受信した部分編集情報に基づいて所望の編集映像に対応する編集情報を編集する編集手段とを備えたという技術的手段を用いる。

【0008】つまり、サーバは、編集情報記憶手段に記憶されている編集情報のうち、どの部分を部分編集情報としてクライアントへ送信するかを設定できるため、各クライアントにより要求する編集情報の種類や内容が異なることにより、サーバにおける編集情報の生成や管理などの処理が複雑化することがない。

【0009】請求項3に記載の発明では、請求項1に記載の映像編集システムにおいて、前記サーバは、さらに、前記編集情報記憶手段に記憶されている編集情報のうち、どの部分を部分編集情報として前記クライアントへ送信するかを設定した設定手段を有し、前記編集情報送信手段は、前記編集情報記憶手段に記憶されている編集情報のうち、前記設定手段により設定された内容に対応した編集情報を部分編集情報として前記クライアントへ送信するという技術的手段を用いる。

【0010】つまり、サーバは、クライアントから送信された送信要求信号により示される内容に対応した編集情報を部分編集情報としてクライアントへ送信することができ、あるいは、自身が設定した内容に対応した編集情報を部分編集情報としてクライアントへ送信することもできる。したがって、クライアントは、サーバに設定されている内容に従う場合は、送信要求信号の種類を状況に応じて変更する手間を省くことができるし、サーバに設定されている内容に対応する編集情報が希望するものでない場合は、送信要求信号を変更することにより、希望する編集情報をクライアントに送信させることができる。

【0011】請求項4に記載の発明では、請求項2また

は請求項3に記載の映像編集システムにおいて、前記編集情報送信手段は、前記受信された送信要求信号の内容に対応して複数種類の部分編集情報を送信可能であり、前記サーバは、さらに、前記複数種類のうち、どの種類を送信するかを前記受信された送信要求信号の内容毎の受信頻度に対応して重み付けする重み付け手段を備えており、前記設定手段は、前記重み付け手段による重み付けに基づいて、設定されている内容を変更可能であるという技術的手段を用いる。

10 【0012】サーバは、受信した送信要求信号の内容毎の受信頻度に対応して重み付けし、その重み付けに基づいて、設定されている内容を変更できる。つまり、クライアントにおいて特定種類の編集情報の使用頻度が高くなり、それに合わせて送信要求信号の内容として同じ内容が設定される頻度が高くなると、サーバは、その内容に対応した編集情報の送信の優先順位を他の種類の編集情報よりも高く設定するため、クライアントは、使用する頻度が高くなつた編集情報を優先的に受信することができるので便利である。

20 【0013】請求項5に記載の発明では、請求項1ないし請求項4のいずれか1つに記載の映像編集システムにおいて、前記編集情報送信手段は、前記部分編集情報の中から、情報量の小さい順に選択して前記クライアントへ送信可能であるという技術的手段を用いる。

【0014】つまり、クリップの開始点および終了点を示すデータなど、必ずしも情報量が多い画像データを送信しなくとも、映像の内容を特定できることが多いため、サーバは、情報量の小さい順に部分編集情報をクライアントへ送信する。したがって、情報量の大きい部分編集情報から送信する場合よりも、通信時間を短縮し、かつ、通信コストを低減できる可能性を高めることができる。

30 【0015】請求項6に記載の発明では、請求項2ないし請求項5のいずれか1つに記載の映像編集システムにおいて、前記クライアントは、さらに、前記設定手段により設定されている内容の変更を要求する変更要求信号を送信する変更要求信号送信手段を備えており、前記サーバは、前記クライアントから送信された変更要求信号を受信する変更要求信号受信手段と、この変更要求信号受信手段が受信した変更要求信号により示される内容に基づいて前記設定手段により設定されている内容を変更する設定内容変更手段とを備えたという技術的手段を用いる。

40 【0016】つまり、クライアントは、サーバにおける設定内容を変更したい場合には、変更要求信号をサーバへ送信することにより、上記設定内容を変更することができる。

【0017】請求項7に記載の発明では、請求項1ないし請求項6のいずれか1つに記載の映像編集システムにおいて、前記サーバは、さらに、映像から編集情報を作

り出す編集情報生成部を備えたという技術的手段を用いる。

【0018】つまり、サーバにおいて編集情報を作り出すため、クライアントにおいて編集情報を作り出す必要がない。したがって、クライアントは、編集情報を作り出すためのコンピュータプログラムを保有する必要がないため、汎用性があり、かつ、編集に必要なコストを低減できる映像編集システムを実現できる。

【0019】請求項8に記載の発明では、請求項7に記載の映像編集システムにおいて、前記編集情報生成部は、前記クライアントから送信された前記送信要求信号を前記送信要求信号受信手段により受信した際に、前記設定手段により設定された内容に対応する編集情報を生成し、その生成した編集情報を前記編集情報記憶手段に記憶するという技術的手段を用いる。

【0020】つまり、サーバは、クライアントから送信要求信号を受信した際に、設定内容に対応する編集情報を生成するため、クライアントが要求していない場合であっても編集情報を生成しておく場合と比較して、編集情報記憶手段における編集情報の記憶領域を効率的に活用できるし、生成および記憶のための無駄な処理を省くことができる。

【0021】請求項9に記載の発明では、請求項1ないし請求項8のいずれか1つに記載の映像編集システムにおいて、前記サーバは、さらに、前記クライアントから送信された編集指示信号を受信する編集指示信号受信手段と、映像を記憶する映像記憶手段と、前記編集指示信号受信手段により受信された編集指示信号にしたがって前記映像記憶手段に記憶されている映像から該当する映像を編集し、編集済映像を前記クライアントへ送信する映像送信手段とを備えており、前記クライアントは、さらに、前記編集手段による編集結果に対応した編集指示信号を前記サーバへ送信する編集指示信号送信手段を備えたという技術的手段を用いる。

【0022】つまり、サーバが映像を記憶し、クライアントからの編集指示にしたがって該当する映像をクライアントへ送信するため、クライアントは、映像編集を行うことができる。したがって、クライアントは、自身に設けられた、あるいは接続された記憶装置に編集映像を記憶することにより、編集映像の記憶された媒体を作成することができる。

【0023】請求項10に記載の発明では、請求項9に記載の映像編集システムにおいて、前記映像送信手段は、前記該当する映像を前記サーバに接続された映像記憶装置へ出力するという技術的手段を用いる。

【0024】つまり、サーバは、該当する映像をクライアントへ送信するのではなく、自身に接続された映像記憶装置へ送信するため、クライアントは、サーバから送信された映像を記憶する必要がない。したがって、クライアントは、サーバから送信された映像を受信するため

の通信コストがかからないとともに、受信した映像を記憶するための記憶領域を確保しておく必要もない。また、クライアントは、上記映像記憶装置に備えられた記憶媒体（たとえば、DVD-R、DVD-RW、ビデオテープなど）をサーバから受け取ることにより、自身が編集した映像を入手できる。

【0025】

【発明の実施の形態】以下、この発明に係る映像編集システムの実施形態について図を参照して説明する。な

10 お、以下に示す各実施形態では、ビデオテープに記録された映像を編集し、その編集映像をDVD-Rに記録する場合を例に挙げて説明する。

【主要構成および主要機能】最初に、この実施形態に係る映像編集システムの主要構成および主要機能について図1および図2を参照して説明する。図1は、映像編集システムの主要構成をブロックで示す説明図であり、図2は、図1に示す映像編集システムの主要機能をブロックで示す説明図である。なお、図2においてブロックで示す各機能は、CPU（図示省略）がコンピュータプログラムを実行することにより実現する機能を示す。また、各記憶部は、ハードディスクやRAMなどの記憶媒体（図示省略）により機能する。

20 【0026】図1に示すように、映像編集システム10は、サーバ20と、サーバ20とネットワーク14を経由して接続されたクライアント40とから構成されている。サーバ20には、ビデオデッキ11およびDVD(Digital Versatile Disc)ライター13が接続されている。DVDライター13は、DVD-Rに映像および音声を記録する機能と、DVD-RまたはDVD-ROMを再生する機能とを有する。サーバ20は、ビデオキャプチャカード21と、記憶部30と、ストリーム編集実行部22と、ネットワークI/F23とを備える。記憶部30は、映像・音声ストリーム記憶部31と、映像を編集するための情報（以下、編集情報と略称する）を記憶する編集情報記憶部32とを備える。また、ネットワーク14は、公衆通信交換網およびインターネットを組み合わせたネットワークである。

30 【0027】ビデオキャプチャカード21は、ビデオデッキ11により再生されたビデオテープ12の映像および音声を読み込み、その読み込まれた映像および音声は、ビデオキャプチャカード21を操作するソフトウェアであるキャプチャカード操作部24により、コンピュータにより処理可能なストリームデータに変換され、映像・音声ストリーム記憶部31に記憶される（図2）。編集情報を作成するソフトウェアである編集情報作成部26は、映像・音声ストリーム記憶部31に記憶されている映像・音声ストリームを用いて編集情報を作成し（図2）、その作成した編集情報を編集情報記憶部32に記憶する。編集情報は、種類毎に所定の条件により重み付けされ、階層化して記憶される（図1）。ストリーム編

集実行部22は、ネットワークI/F23を介してクライアント40から受信した編集指示にしたがって映像・音声ストリーム記憶部31に記憶されている映像・音声ストリームを編集し、その編集された映像・音声ストリームをDVDライター13によりDVD-Rに記録する。

【0028】クライアント40は、パーソナルコンピュータ(以下、PCと略称する)42と、ディスプレイ43と、キーボード44と、マウス45と、ネットワークI/F41とを備える(図1)。クライアント40は、データの送受信を行うソフトウェアであるデータ送受信部50(図2)により、サーバ20に対して編集情報の送信要求を示す信号(以下、送信要求信号と略称する)を送信し、その送信要求信号により示される内容に対応した編集情報をサーバ20から受信する。この受信された編集情報は、編集情報の解読を行うソフトウェアである編集情報解読部48(図2)により解読され、その解読結果は、ディスプレイ43に編集画面を表示するソフトウェアである編集画面表示部46により、ディスプレイ43に表示される。また、上記解読結果は、編集内容記憶部49(図2)にも記憶される。

【0029】編集画面操作部47は、ディスプレイ43に表示された編集画面上で編集作業を行うためのソフトウェアであり、公知のソフトウェアである。たとえば、その編集画面の表示形態は、たとえば米国のマイクロソフト社が開発したOSであるWindows(Windowsは、マイクロソフト社の登録商標)方式の表示形態である。この表示形態によれば、マウス45によりポインタを画面上を移動させ、クリック操作するだけの簡単な操作により、編集作業を行うことができる。その編集結果は、編集内容記憶部49(図2)に記憶され、編集指示作成部49(図2)は、編集内容記憶部49に記憶されている編集内容に対応した編集を行うようにサーバ20に対して指示するための信号(以下、編集指示信号と略称する)を作成する。そして、サーバ20のストリーム編集実行部22(図1)は、クライアント40から送信された編集指示信号に従って映像・音声ストリーム記憶部31から映像・音声ストリームを読み出して編集を実行する。

【0030】【編集情報の種類】次に、編集情報の種類について図3を参照して説明する。図3は、編集情報記憶部32に記憶された編集情報の種類を示す説明図である。

#### (1) サムネイル画像32a

サムネイル画像とは、元の画像の大きさを縮小したものである。サムネイル画像は、周知のアルゴリズムにより生成可能であり、たとえば元画像を低域フィルタを通して高周波成分を取り除いた後、画素を間引くことで画像を縮小できる。MPEGやJPEGのI-frameであれば、DC成分を抽出することで容易にサムネイルを

作成することができる。このため、画質(解像度)は元画像よりも劣るが、映像の1コマを見ることができるため、映像の内容を十分に把握することができる。

【0031】編集作業として、クリップ単位でのカットや並べ替えを行うときは、クリップの繋ぎ目、つまりクリップの先頭および末尾付近のフレームを見て判断するため、そのフレームのサムネイル画像を編集情報としてサーバからクライアントへ送信するのが望ましい。また、通常のビデオ映像は、30フレーム/秒であり、連続した2つのフレームは、シーンの変わり目(クリップの境界)を除くと、あまり差がない。したがって、同じ数のフレームのサムネイルを送信することを考えた場合、連続したフレームよりも、ある間隔(たとえば1秒)置きにフレームを送信した方が、クリップ全体を把握するのに役立つ。また、データのサイズを小さくすることができる。たとえば、元画像が解像度720×480RGB形式(BMP)のデータ量は、1,036,854バイトであるが、720×480JPEG圧縮したもので67,289バイトであり、元画像のデータ量の約6.5%となる。また、元画像の解像度を変え、360×240としたJPEG圧縮では21,478バイトであり、元画像のデータ量の約2.1%となる。さらに、180×120JPEG圧縮では7,399バイトであり、元画像のデータ量の約0.7%となる。

【0032】(2) 2値画像32b  
フルカラーの画像ではなく、白黒の2値画像をサムネイルとして用いても、その画面を十分に認識することができるため、映像の内容を十分に把握することができる。単純なビットマップ画像で考えた場合、フルカラーでは、1ドットに対して24ビットのデータが必要になるが、2値画像では、1ビットで済む。したがって、サムネイルのデータ量を1/24に減らすことができる。また、2値画像データを圧縮することにより、さらにデータ量を減らすことができる。2値画像は、各画素について、その値が、あるしきい値より大きい場合には「1」、値が小さい場合には「0」を割り当てることにより生成できる。たとえば、カラー画像の場合は、まず、グレースケールに変換する。RGBカラー画像をグレースケールに変換するには、各画素について以下の変換式を作用させる。

【0033】 $Y=0.299R+0.587G+0.144B$   
次に、上記グレースケールに対して、しきい値を設定するが、良好な(視認性の良い)2値画像を生成するためには、上記しきい値の求め方が重要である。元々の画像が2値のもの、たとえば白い紙に黒いインクで文字が書かれているものなどは、輝度ヒストグラムには、2つのピークを生じることになるため、その中間点にしきい値を設定すれば良い。一方、元々の画像が2値でない画像は、最適なしきい値を求めることが難しいが、単純に、対象とする画像の最大値および最小値の中間の値をしき

い値として2値化する。

【0034】(3) 画像の一部32c

画像の一部をトリミングしたデータをサムネイルとして用いても、その画面を十分に認識することができるため、映像の内容を十分に把握することができる。つまり、画像の一部に、その画像の内容を特徴付ける部分が存在する場合は、その部分させ見ることができれば、その画像が含まれているクリップの内容を知ることができる。図4は、画像の一部をトリミングする様子を示す説明図である。同図に示すように、横n(たとえば80画素)×縦m(たとえば60画素)により囲まれた矩形Aを取り出して新たなビットマップ画像Bを生成する。このとき、画像Bの解像度は画像Aの数分の1(たとえば1/4)となるが、その画像Bには、画像Aの内容を特徴付ける部分が含まれているため、画像Aと比べて遜色ない。しかも、データ量を数分の1(たとえば1/4)に減らすことができる。画像の一部32qは、図4に示すように、トリミングにより切り出したい矩形Aの情報(a, b, m, n)を与えることにより、元画像の内で、矩形Aの範囲に該当する画素のみを抽出することにより生成できる。

【0035】(4) 線画32d

フルカラーの画像あるいは2値画像ではなく、線画をサムネイルとして用いても、その画面を認識することができるため、映像の内容を十分に把握することができる。線画は、2値画像の輪郭を取出したものであるため、2値画像よりもデータ量を大幅に削減できる。また、線画を生成できる部分が少ない(まばらな)場合には、チーン符号により極めて少ないデータ量で画像を表現できる。また、線画は、画像の2値化、辺縁の検出、細線化の3ステップにより生成できる。

【0036】(5) クローズド・キャプション(CC)  
32e

クローズド・キャプションは、ビデオ信号の21番目の走査線に埋め込まれている。そのクローズド・キャプションとして、クリップの字幕を示すデータを用いることにより、クリップ中のシーンを特定できる。既存のMPEGエンコーダには、ビデオのキャプチャ装置によりその部分を取り出し、MPEGストリームに含める機能を持つものがある。また、通常用いられているMPEGデコーダは、上記部分を取出す機能を備えているので、それによりデータを抽出することで、クローズド・キャプションを取出すことができる。

【0037】(6) テープラベル情報32f

ビデオテープの外見から読み取れる情報をテープラベル情報として用いることにより、複数のビデオテープがある場合にテープを識別できる。たとえば、次の情報をテキストデータにする。形式(MiniDV, Hi8, Video8, VHS, VHS-Cなど)(8バイト以内)、メーカー名(20バイト以内)、品名(20バ

ト以内)、テープ長[分](4バイト以内)、属性(MP、MEなど)(8バイト以内)、録画方法(SP、LP、標準、3倍など)(8バイト以内)、ラベルシート内容(256バイト以内)。また、各行には、改行コード(2バイト)が入る。テープラベル情報32fのデータサイズは、テープ1本のデータは最大338バイトであり、ビデオテープが10本あるとしても合計3380バイトである。これを64kbpsの通信帯域では、約0.4秒で送信完了するため、データ量が小さいと言える。テープラベル情報32fは、サーバ側の作業者が、ビデオテープの外見から読み取り、その読み取った情報をテキストエディタなどにより入力してテキストデータを作成する。

【0038】(7) 音情報32g

①音声ストリームをMPEG-Audioにエンコードしたものの先頭のある時間分を分割して保持する。たとえば約3秒間分を分割して保持する場合は、オーディオフレームで割り切れ、かつ最も3秒に近い時間分を分割して保持する。先頭のある時間分の音を聞くことにより、映像の内容を十分に把握することができる。

②クリップの一部あるいは全体の平均の音量を示したり、最小および最大音量を示したりすることで、そのクリップの内容を類推することができる。たとえば、山間の静かなところで景色を撮影した内容のクリップの後に自動車レースを撮影したクリップが記録されているような場合は、クリップ間で平均音量または最大音量がかなり異なるため、それらの情報を知ることにより、クリップの内容を類推できる。

【0039】各クリップ毎の平均音量を示す場合は、1

30 6ビットサンプリングで音量を示すとすると2バイト、8ビットサンプリングで示すと1バイトのデータサイズとなる。ここに挙げたサイズは、クリップの長さに関わらず一定であり、データ量は非常に小さい。ただし、単純に音量を平均化するだけでは、クリップの長さが長くなるほど特徴を拾い出すのが困難になるため、クリップのごく一部だけを取出すようにする。各クリップの音量情報32sとしては、種々の情報が考えられるが、たとえばクリップ内の全サンプルの音量を平均化してそのクリップの音量の代表値とする。また、クリップの先頭n秒(たとえば3秒)分のサンプルの音量を平均化して

40 (あるいは最大値をとり)、そのクリップの音量の代表値とすることもできる。さらに、クリップ内でサンプルをn秒(たとえば5秒)ずつに分け、分けた部分についてそれぞれを平均化して(あるいは最大値をとり)、クリップ内での音量の分布を表すようにしても良い。この場合、データ量は、クリップの長さ(秒)/5倍となる。さらに、先頭オーディオフレームの帯域の周波数分割情報を作成しても良い。

【0040】③音のスペクトル

50 図5は、音の9個の帯域のスペクトルを示すグラフであ

る。横軸が周波数であり、縦軸が音のエネルギーである。このグラフによれば、音の周波数特性（=音の高低）を視覚化することができ、音色を視覚的に把握することができる。データの大きさは、1サンプルを4バイトで表すとすると、サンプリング周波数32kHzのPCMデータ1秒分は、 $4\text{バイト} \times 32,000 = 128,000$ バイトとなるが、音のスペクトルの1帯域分を4バイトで表すとすると、 $4\text{バイト} \times 9 = 36$ バイトで済むため、データ量を大幅に削減することができる。音のスペクトルは、たとえばフーリエ変換やウェーブレット変換などの周知の手法により求めることができる。

【0041】(8) クリップの開始点・終了点32h クリップ単位の編集作業（たとえば、挿入、削除、複写、並べ替えなど）を行う場合において、素材となるビデオ映像の内容を時間単位で把握できる場合は、各クリップの開始点および終了点が分かれば、各クリップの長さを求めることができるため、その求めた長さに基づいてクリップの内容をそれぞれ特定できるので、クリップ単位の編集作業を行うことができる。図6は、クリップの開始点・終了点を示すデータ構造の説明図である。1つの開始点または終了点につき、時分秒をそれぞれ1バイトずつで表し、フレーム数を1バイトで表すため、計4バイトのデータ量である。たとえば、素材となるビデオ映像の全体が100個のクリップで構成されている場合でも、開始点・終了点を示すデータ量の合計は4バイト $\times 100 + 1$ バイト = 404バイトであり、データ量は極めて小さい。

【0042】ここで、クリップの開始点・終了点を検出する手法の一例を図7を参照して説明する。図7は、サーバ20が編集情報としてのクリップの開始点・終了点を検出する際の処理の流れを示すフローチャートである。この手法は、予めクリップの切れ目としたい箇所にブランク画像（真っ黒な画面や真っ白な画面など）を10フレーム連続で挿入しておくことを特徴とする。サーバ20は、素材となるビデオ映像を構成するフレーム画像を取得し（ステップ（以下、Sと略す）12）、ブランク画像を検出すると（S14: Yes）、ブランク画像をカウントするカウンタのカウンタ値に「+1」を加算する（S16）。続いてサーバ20は、カウンタのカウンタ値が「10」に達するまでフレーム画像の取得からブランク画像の検出およびカウンタ値の加算を繰り返し（S12～S16）、カウンタ値が「10」に達すると（S18: Yes）、その10フレーム目のブランク画像に対応する時間をクリップの切れ目として設定する（S20）。

【0043】続いてサーバ20は、ブランク画像のカウンタをゼロリセットし（S22）、次のフレーム画像が存在する場合は（S26: Yes）、再びS12～S22を繰り返す。また、サーバ20は取得したフレーム画像がブランク画像でない場合は（S14: No）、ブ

ンク画像のカウンタをゼロリセットし（S24）、次のブラック画像を取得するまで待機する（S26、S12～S14）。また、クリップの開始点・終了点を検出する他の手法として、①録画日時の不連続点を検出する手法（たとえば特開平11-46336号公報）。②素材のビデオ映像を先頭から一定時間毎に分割し、それらの1つ1つを自動的にクリップとして設定する手法。これは、経過時間毎にクリップ分けしたい場合に有効な手法である。③素材のビデオ映像の中に含まれるフェードインまたはフェードアウトの箇所を検出し、その検出された箇所をクリップの切れ目に設定する手法。これは、通常、フェードインまたはフェードアウトは、場面の変わり目に行うため、場面の変わり目毎にクリップ分けしたい場合に有効な手法である。この場合、前述のブランク画像の検出と同じアルゴリズムを適用できる。④サーバ側の担当者がクライアントの希望に従って素材のビデオ映像を見ながら1つ1つクリップを設定して行く手法。これは、手間はかかるが、クライアントの意図しないクリップ分けになることが少なくなるという利点がある。また、他の手法で大雑把にクリップ分けした後にこの手法を用いると、クリップ分けを効率良く行うことができる。

【0044】(9) クリップ数・番号32i クリップ数は、クリップ開始点の数または終了点の数に対応するため、前述のクリップ開始点・終了点を検出するアルゴリズムを実行することにより求められる。また、クリップ番号は、クリップ開始点を検出したときに付すことができる。最初に大雑把にクリップ数またはクリップ番号を指定し、その後徐々に絞り込んで行くことにより、希望するクリップを見付けることができる。

【0045】(10) フレーム数32j クリップ内のフレーム数から、そのクリップの長さを類推できるため、クリップの長さにより、内容を把握できる場合に有効な手法である。また、記録した順序で送信することにより、経過時間を知ることもできる。また、2バイトで表現できるフレーム数は、 $2^{16} - 1 = 65,535$ フレームであり、これは、フレームレートが30fpsであるとすると、約36分間の長さに相当する（ $65,535 / 30 = \text{約}36$ ）。したがって、総てのクリップが約36分以内であれば、全体が100個のクリップを表現するために必要なデータ量は、 $2\text{バイト} \times 100 = 200$ バイトとなり、非常に小さなデータ量で済む。フレーム数は、各クリップ内のフレームの数をカウントアップして生成する。また、クリップ開始点・終了点からクリップの時間を求め、その時間にフレームレートを乗算することにより生成できる。

【0046】(11) シーンチェンジ点数32k 素材のビデオ映像の中において、どのくらいのシーンチェンジ点があるかを把握している場合は、シーンチェンジ点数を指定することにより、編集作業を行うことができる。

きる。

【0047】(12) 動きベクトルm

画面のアクティビティが高い部分は、動きが多いので、編集上、重要な場合であることが予測できる。また、アクティビティが極端に低い場面は、静止画であると考えられるが、ビデオ撮影に失敗した部分であることも多い。そこで、画面のアクティビティの高低を検出することにより、編集点を見付ける手法を用いることができる。DVDやデジタルカメラなどで適用されているMPEG(Motion Picture Experts Group)などの画像圧縮方式では、フレーム間の動きベクトルを用いて圧縮を行っているため、その動きベクトル情報を用いて画面のアクティビティを表すことができるので、そのアクティビティを画像の識別に用いることができる。

【0048】図8(A)は、画面の動きを示す矢印の説明図であり、図8(B)は、動きベクトルの座標の一部を示す説明図である。たとえば、編集画面においてサムネイルを表示すべき部分に、図8(A)に示すように矢印を表示する。これにより、画面がどちらの方向に、どの程度動いているかを示すことができる。たとえば、左側の矢印により、画面が右上に動いていることが分かり、その矢印よりも短い右側の矢印により、画面が奥行き方向に速度を落として動いていることが分かる。また、アクティビティの高低を、赤色から青色までのグラデーション(濃淡の段階的変化)で表し、それを矢印の背景として用いることにより、映像のアクティビティを感覚的に表すこともできる。

【0049】データ量は、MPEGの符号化単位であるGOP(Group of Pictures)毎に、数十バイトで表現できるので、極めて小さい。また、動きベクトルは、MPEG圧縮されたストリームから取り出すことにより生成する。MPEGストリームのフォーマットは、ISO規格の13818-2により規定されており、既存のMPEGデコーダと同等の処理で動きベクトルを取り出すことができる。MPEGにおいては、PBピクチャのPBマクロブロックに動きベクトルが用いられる。Iピクチャ(およびタイプIのマクロブロック)には動きベクトルは存在しない。また、BピクチャのBタイプマクロブロックには、後ろ向きの動きベクトルも存在するが、ここでは、それを用いず、順方向のみの動きベクトルを使用する。このように、総てのフレームの総てのマクロブロックに動くベクトルが存在するわけではないので、動きベクトルに関する処理は、ある単位長のデータについて計算し、それを平均化して用いる。たとえば、単位長は、MPEGのGOPを用い、それを15フレーム(30fpsにおいて0.5秒)とする。

【0050】動きベクトルは、図8(B)に示すように、各マクロブロック毎の1組の座標として得られる。これらのベクトル値から単位長の映像における動きベクトルの代表値を取出す手法を以下に示す。

①動きベクトルの和を求める。各動きベクトルの和は、動きベクトルの存在するマクロブロックの数で正規化するので、平均値の計算を行って各動きベクトルの和を求める。これにより、画面全体がどの方向に動いているかを示す代表的な動きベクトルを求めることができる。

②動きベクトルの2乗和を求める。単純な和では、画面の平行移動は検出できるが、たとえば回転の場合、単純な和では結果が「0」になって動きを検出できない。このように、単純な方向を取り出せない場合、大きさのみでアクティビティを求める。そのためには、各ベクトルの長さの2乗和をブロック数で正規化して用いる。③動きベクトルの最大値を取出す。映像の一部のみに動きがある場合、意味のある動きベクトルを持つブロックはさほど多くない。そのような場合には、ブロック数で正規化すると、特徴的な値を取出すことができない。そこで、意味のある動きベクトルとして、最大の大きさを持つ動きベクトルを代表値として採用すれば、上記の場合に対応できる。また、これと同様の手法として、ブロック数の20%までをベクトルの大きさの大きいものから選択するという手法もある。さらに、この方法と、①および②の手法を組み合わせて用い、総てのブロックのベクトルではなく、ベクトル値の大きいものに対して、和や2乗和を求めてよい。

【0051】(13) 単色画像32n

ある画面の中で最も多く使われている色や、全体を平均化した色を代表色として画像の代わりに表示する。たとえば、サムネイルを並べて表示する際、転送されていない部分を単なるブランクで表示するのではなく、画面に関連のある色で表示することにより、映像の雰囲気が分かる。特に、シーンの変わり目は単色画像のみでも識別可能である。また、単色であれば、色の値をHTML(Hyper Text Markup Language)中に記述するだけでクライアント側のブラウザで表現できる。このとき、HTMLに記述する際は、#12A844のようにRGB形式で記述するだけでブラウザに表示させることができるため、色データをそのまま送信する場合と比較してデータサイズは非常に小さい。具体的には、<table width="20" height="20" bgcolor="#FF0000"><tr><td></td></tr></table>の計74バイトでブラウザ内に20×20ピクセルサイズの赤い四角を表示させることができる。たとえば、画像をRGB成分で解析し、RGBの分布から最頻値または平均値を求め、単色を決定する。

【0052】[主な処理内容] 次に、サーバ20およびクライアント40が実行する主な処理内容を説明する。

(重み付け) 最初に、サーバ20が編集情報をクライアント40へ送信する際に、どの編集情報を優先して送信するかの順位を重み付けする手法について説明する。サーバ20の編集情報記憶部32には、図3に示したように各種の編集情報が記憶されているが、送信時の優先順位は、図9に示す優先順位テーブル32pにより決定す

る。優先順位テーブル32pは、各クライアント毎に設定されており、ユーザID（図9に示す例ではUSER000001）と、優先順位と、クライアントからの要求頻度と、クライアントへの送信回数とを対応付けて構成される。図9に示す例では、優先順位1位がサムネイルであり、その要求頻度は38であり、送信回数は87となっている。優先順位は、要求頻度および送信回数に基づく学習により入れ替わる。

【0053】（サーバ20およびクライアント40が実行する主な処理の流れ）次に、サーバ20およびクライアント40が実行する主な処理の流れについて図11および図12を参照して説明する。図11はサーバ20が実行する主な処理の流れを示すフローチャートであり、図12はクライアント40が実行する主な処理の流れを示すフローチャートである。サーバ20は、ビデオデッキ11（図1）により再生されたビデオテープ12の映像に基づいて前述した各種の編集情報を生成し、編集情報記憶部32（図2）に記憶する（図11のS10）。また、クライアント40が、編集情報の送信要求を示す信号（以下、送信要求信号を称する）をサーバ20へ送信すると（図12のS500）、サーバ20は、その送信要求信号を受信し（図11のS50）、その送信要求信号により示される内容に対応した編集情報（以下、部分編集情報と称する）を編集情報記憶部32から読み出してクライアント40へ送信する（S100）。

【0054】ここで、サーバ20が図11のS100において実行する部分編集情報送信処理の流れについて、それを示す図14のフローチャートを参照して説明する。サーバ20は、クライアント40から送信要求信号を受信したことと示す送信要求フラグをチェックし（S102）、フラグがある場合は（S104：Yes）、送信要求信号により示される内容に対応した部分編集情報を編集情報記憶部32から読み出し、その読み出した部分編集情報をクライアント40へ送信する（S106）。続いてサーバ20は、送信要求フラグをリセットし（S108）、送信要求フラグをチェックするが（S102）、S108においてリセットされているため送信要求フラグなしと判定する（S104：No）。続いてサーバ20は、部分編集情報を送信中でない、つまり通信回線が空いていると判定すると（S110：No）、部分編集情報を送信する（S112）。

【0055】なお、送信する部分編集情報がサムネイルである場合は、前述のようにサムネイルは、箇所により有用性が異なるので、1つのクリップ中では、①クリップの先頭、末尾のサムネイル。②ある時間（たとえば1秒）単位のサムネイル（これも先頭、末尾から中央に向かって送信する）。③他のサムネイル。また、クリップの送信順序は、自動では先頭から送信するが、クライアントの指示がある場合は、クライアントが最後に要求したクリップの優先順位を高くする。そしてサーバ20

は、部分編集情報の終了を待たずに送信要求フラグをチェックする（S102）。つまり、クライアントからの送信要求に速やかに応答し、その送信要求に対応した部分編集情報を送信する。仮に、クライアントからの送信要求を検出した時点で、前記順位による方法に従って送信中の情報がある場合には、その送信をキャンセルしてクライアントからの指示を優先させる。

【0056】そして、クライアント40は、サーバ20から送信された部分編集情報を受信し（図12のS550）、その受信した部分編集情報をディスプレイ43に表示し、キーボード44およびマウス45を用いて編集作業を行う（S600）。そして、クライアント40は、編集作業を終了すると、編集指示信号をサーバ20へ送信する（S650）。続いて、サーバ20は、クライアント40から送信された編集指示信号を受信し（図11のS150）、ストリーム編集部22（図2）は、受信した編集指示信号に基づいて、映像・音声ストリーム記憶部31から映像・音声ストリームを読み出し、その編集を実行し、その編集結果をDVDライタ13へ出力する（S200）。これにより、DVDライタ13は、入力した編集結果をDVD-Rに書き込み、クライアント40がビデオテープ12を編集した内容が記録されたDVD-Rが完成する。

【0057】また、サーバ20は、クライアント40からの送信要求の頻度に対応して優先順位の設定内容を変更する（図11のS250）。ここで、サーバ20が実行する設定内容変更処理の流れについて、それを示す図13のフローチャートを参照して説明する。サーバ20は、クライアント40から送信された送信要求信号を受信すると（S252：Yes）、その受信した送信要求信号により指定された編集情報に対応付けられている要求頻度（図9）に所定値を加算し（S254）、ユーザ情報を更新する（S256）。

【0058】続いてサーバ20は、所定値を加算した後の要求頻度と、自身よりも優先順位の高い他の編集情報の要求頻度とを比較し（S258）、上記要求された編集情報の要求頻度よりも低い要求頻度の他の編集情報が存在する場合は（S260：Yes）、上記要求された編集情報と上記他の編集情報との優先順位を入れ替える（S262）。つまり、クライアントの要求頻度を学習により優先順位に反映させることができる。なお、ビデオテープ1本に対して1つのみ存在する編集情報から、クリップ1つに対して1つのみ存在する編集情報、あるいは各フレーム毎に1つ存在する編集情報まで存在するため、上記所定値は、編集情報により異なる。たとえば、単色画像、2値画像および線画については、それぞれ所定値は「1」であり、音量および動きベクトルは、それぞれ「10」であり、クリップ開始点・終了点およびクリップ数・番号については、それぞれ「50」とする。

【0059】また、あるクリップの編集情報（たとえばサムネイル）を送信中に他のクリップの同種の編集情報がクライアントから要求された場合は、次の送信時には、前に送信していた編集情報ではなく新たに指定された編集情報を含むクリップの送信を開始する。さらに、あるクリップの編集情報（たとえばサムネイル）を送信中に、クライアントの編集作業により、その編集情報がカットされた場合に、そのクリップの情報はもう必要ないと考えられるので、送信を停止し、他の編集情報を送信することができる。また、そのクリップが編集結果に採用されることが確定した場合は、そのクリップの情報が必要になる可能性は小さいので、その時点で、編集情報の送信を停止することができる。なお、総てのフレームに対するサムネイルは、ビデオ映像を送信することと等しいため、優先順位から外すことができる。

【0060】[実施形態の効果]

(1) 以上のように、上記実施形態の映像編集システム10を使用すれば、クライアント40は、映像を編集するために必要な編集情報をサーバ20から受信し、その編集情報により編集を行うことができるため、総ての映像を受信する必要がない。したがって、従来のように、クライアント40が総ての映像を受信してから編集作業を行う手法よりも、編集作業を行うことを決定してから実際に編集作業を開始するまでにかかる時間を短縮することができる。また、サーバ・クライアント間の通信時間を短縮できるため、通信コストを低減することもできる。

(2) また、サーバ20は、クライアント40から送信された送信要求信号により示される内容に対応した編集情報をクライアント40へ送信することができる。したがって、クライアント40は、サーバ20から受信した編集情報が希望するものでない場合は、送信要求信号を変更することにより、希望する編集情報を受信することができる。また、クライアント40は、サーバ20に設定されている内容に従う場合は、送信要求信号の種類を状況に応じて変更する手間を省くことができる。

【0061】(3) さらに、サーバ20は、編集情報の送信の優先順位を送信要求頻度に対応して重み付けし、その重み付けに基づいて、設定されている優先順位を変更できる。つまり、クライアント40において特定種類の編集情報の使用頻度が高くなると、サーバ20は、その編集情報の送信の優先順位を他の種類の編集情報よりも高く設定するため、クライアント40は、使用する頻度が高くなった編集情報を優先的に受信することができるので便利である。

(4) またさらに、サーバ20において編集情報を作り出すため、クライアント40において編集情報を作り出す必要がない。したがって、クライアント40は、編集情報を作り出すためのコンピュータプログラムを保有する必要がないため、汎用性を高めることができ、かつ、

編集に必要なコストを低減できる映像編集システムを実現できる。

【0062】(5) そしてさらに、サーバ20が映像・音声ストリームを記憶し、クライアント40からの編集指示にしたがって該当する編集情報をクライアント40へ送信するため、サーバ20は映像を記憶しておく必要がない。したがって、サーバ20は映像・音声ストリームを記憶するための負担を軽減することができる。

(6) さらに、サーバ20は、編集した映像をクライアント40へ送信するのではなく、自身に接続されたDVDライタ13へ出力するため、クライアント40は、サーバ20から送信された映像を記憶する必要がない。したがって、クライアント40は、サーバ20から送信された映像を受信するための通信コストがかからないとともに、受信した映像を記憶するための記憶領域を確保しておく必要もない。また、クライアント40は、サーバ20からDVD-Rを受け取ることにより、自身が編集した映像を入手できる。

【0063】[他の実施形態]

20 (1) 編集情報の組み合わせ

優先順位付けされる編集情報としては、上述した単独ではなく、以下に示すように組み合わせて用いることができる。この場合、特に組み合わせにより有用性が高くなるが、データ量がそれほど増加しない組み合わせを用いると効果的である。

① 1つのクリップにおけるフェードイン点およびフェードアウト点と、それらに対応するサムネイル2枚  
この組み合わせによれば、クリップの内容の把握が容易になる。

30 ② クリップ先頭のサムネイルおよび動きベクトル  
動きベクトルのデータだけでは、有用性は低いが、クリップ先頭のサムネイルを動きベクトルのデータに付加することで、サムネイルおよび動きから画像を推測できるので、有用性を高めることができる。

【0064】③ 音情報および先頭のサムネイル

画像よりも音の方がデータ量が少ないので、画像は1枚のみとし、他は音情報だけを送信することで、データ量を減少させることができる。

④ 単色画像および線画

40 単なる線画に、その画面を代表する色を付けることにより、クリップの内容を感覚的に分かり易くすることができる。2値画像を白黒ではなく、代表色により表示しても、同様の効果が得られる。

⑤ サムネイルおよび画像の一部

画像の一部を送信する場合、毎回送信する領域を指定していくには手間がかかるため、サムネイルとする領域を一律に固定するのが自然である。なお、上記のように編集情報を組み合わせる場合には、以前にどのような編集情報が送信されているかにより、次に送信すべき編集情報の優先順位付けを変えることが効果的である。たとえ

ば、クリップの先頭のサムネイルは、極めて有用性が高いので、まずクリップの先頭のサムネイルを送信し、その後、動きベクトルや単色画像、音情報などを送信するように制御することが望ましい。

【0065】(2) クライアント40において上記優先順位を直接変更することもできる。図10は、クライアント40に設けられたディスプレイ43(図1)の表示内容を示す説明図である。図10に示すように、ディスプレイ43に、編集情報の優先順位の学習状況を表示する。図10に示す例では、優先順位が表形式で表示され、優先順位1位がフレーム数であり、2位がサムネイルとなっている。また、表の下部には、上記学習をクリヤーするか、保存するかを選択するチェックボックスと、選択を確定するボタン(g o ! !)とが表示されている。優先順位の変更は、文字ボックス内に文字を入力する、優先順位を示す数字を変更する、文字ボックスをドラッグして移動させるなどの手法により行うことができる。また、マウス45の操作によりクリヤーをチェックし、g o ! !を押すと、表示されている学習状況の内容が消去される。また、保存をチェックし、g o ! !を押すと、表示されている学習状況の内容が保存される。なお、クライアント40において、優先順位の変更または消去の操作が行われた場合には、その操作に対応する変更要求信号がサーバ20へ送信される。つまり、クライアント40は、サーバ20における設定内容を変更したい場合には、変更要求信号をサーバ20へ送信することにより、上記設定内容を変更することができる。なお、この構成を採用した場合の映像編集システムが、請求項6に係る映像編集システムに対応する。

【0066】(3) 編集情報のデータ量の大小により、編集情報の送信の優先順位付けを行い、編集情報記憶部32に記憶しておくこともできる。図15は、データ量の小さい編集情報を優先順位を高く設定した場合の説明図である。つまり、フレーム数、クリップの開始点・終了点を示すデータなど、必ずしも情報量が多い画像情報などを送信しなくとも、映像の内容を特定できる場合があるため、サーバ20は、情報量の小さい順に編集情報をクライアント40へ送信する。したがって、情報量の大きい編集情報から送信する場合よりも、通信時間を短縮し、かつ、通信コストを低減できる可能性を高めることができる。なお、この構成を採用した場合の映像編集システムが、請求項5に係る映像編集システムに対応する。

【0067】(4) サーバ20の編集情報生成部26は、クライアント40から送信された送信要求信号を受信した際に、設定内容に対応する編集情報を生成し、その生成した編集情報を編集情報記憶部32に記憶するように構成することもできる。この構成によれば、サーバ20は、クライアント40から送信要求信号を受信した際に、設定内容に対応する編集情報を生成すればよいた

め、クライアント40が要求していない場合であっても編集情報を生成しておく場合と比較して、編集情報記憶部32における編集情報の記憶領域を効率的に活用できるし、生成および記憶のための無駄な処理を省くことができる。なお、この構成を採用した場合の映像編集システムが、請求項8に係る映像編集システムに対応する。

【0068】(5) 前述の実施形態では、サーバ20から編集映像および音声をサーバ20に接続されたDVDライタ13に出力する場合を説明したが、サーバ20から編集映像および音声をクライアント40へ送信し、クライアント40に備えられた記憶装置あるいはクライアント40に接続されたDVDライタなどの映像記憶装置に書き込むように構成することもできる。この構成によれば、編集映像および音声が記録された記録媒体をクライアント40において作成することができる。なお、この構成を採用した場合の映像編集システムが、請求項9に係る映像編集システムに対応する。

(6) 前述の実施形態では、この発明に係る映像編集システムとして、1台のクライアント40と1台のサーバ

20 20とがネットワーク14を経由して接続されている場合を例に挙げて説明したが、複数のクライアントと1台のサーバ、または1台のクライアントと複数のサーバ、あるいは複数のクライアントと複数のサーバの組み合わせに対してもこの発明を適用できることは勿論である。

【0069】【各請求項と実施形態との対応関係】サーバ20の編集情報記憶部32が、請求項1に係る編集情報記憶手段に対応し、データ送受信部27が、送信要求信号受信手段および編集情報送信手段に対応する。また、クライアント40のデータ送受信部50が、請求項

30 1に係る送信要求信号送信手段および編集情報受信手段に対応し、編集画面表示部46、編集画面操作部47、編集情報解読部48および編集内容記憶部49が、編集手段に対応する。また、優先順位テーブル32p(図9)が、請求項2に係る設定手段に対応し、サーバ20の編集情報生成部26が、請求項7に係る編集情報生成部に対応する。さらに、サーバ20のデータ送受信部27が、請求項9に係る編集指示信号受信手段に対応し、クライアント40のデータ送受信部50が、編集指示信号送信手段に対応する。またさらに、サーバ20の映像

40 ・音声ストリーム記憶部31が、請求項9に係る映像記憶手段に対応し、DVDライタ13が、請求項10に係る映像記憶装置に対応する。

【0070】そして、サーバ20が実行するS50(図11)が、請求項1に係る送信要求信号受信手段として機能し、S100が編集情報送信手段として機能する。また、クライアントが実行するS500(図12)が、請求項1に係る送信要求信号送信手段として機能し、S550が編集情報受信手段として機能し、S600が編集手段として機能する。さらに、サーバ20が実行するS252～S262(図13)が、請求項4に係る重み

付け手段として機能し、S150(図11)が請求項9に係る編集指示信号受信手段として機能し、S200が映像送信手段として機能する。またさらに、クライアント40が実行するS650(図12)が、編集指示信号送信手段として機能する。

【図面の簡単な説明】

【図1】この発明の実施形態に係る映像編集システムの主要構成をブロックで示す説明図である。

【図2】図1に示す映像編集システムの主要機能をブロックで示す説明図である。

【図3】編集情報記憶部32に記憶された編集情報の種類を示す説明図である。

【図4】画像の一部をトリミングする様子を示す説明図である。

【図5】音の9個の帯域のスペクトルを示すグラフである。

【図6】クリップの開始点・終了点を示すデータ構造の説明図である。

【図7】サーバ20が編集情報としてのクリップの開始点・終了点を検出する際の処理の流れを示すフローチャートである。

【図8】図8(A)は、画面の動きを示す矢印の説明図であり、図8(B)は、動きベクトルの座標の一部を示す説明図である。

【図9】優先順位テーブル32pの構成を示す説明図である。

【図10】クライアント40に設けられたディスプレイ\*

\* 43の表示内容を示す説明図である。

【図11】サーバ20が実行する主な処理の流れを示すフローチャートである。

【図12】クライアント40が実行する主な処理の流れを示すフローチャートである。

【図13】サーバ20が実行する設定内容変更処理の流れを示すフローチャートである。

【図14】サーバ20が実行する編集情報送信処理の流れを示すフローチャートである。

10 10 【図15】データ量の小さい編集情報を優先順位を高く設定した場合の説明図である。

【図16】従来から行われている動画の編集手法の一例を示す説明図である。

【図17】従来から行われている動画の編集手法の一例を示す説明図である。

【図18】従来から行われている動画の編集手法の一例を示す説明図である。

【符号の説明】

10 10 10 映像編集システム

11 11 11 ビデオデッキ

13 13 13 DVDライタ(映像記憶装置)

20 20 20 サーバ

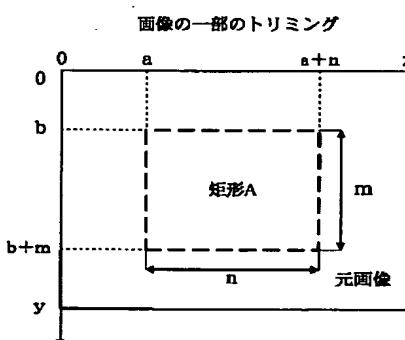
31 31 31 映像・音声ストリーム記憶部(映像記憶手段)

32 32 32 編集情報記憶部(編集情報記憶手段)

32p 32p 32p 優先順位テーブル(設定手段)

40 40 40 クライアント

【図4】

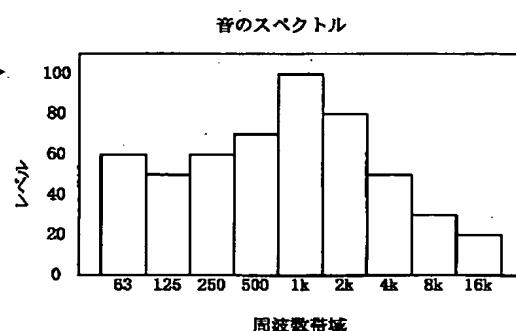


【図6】

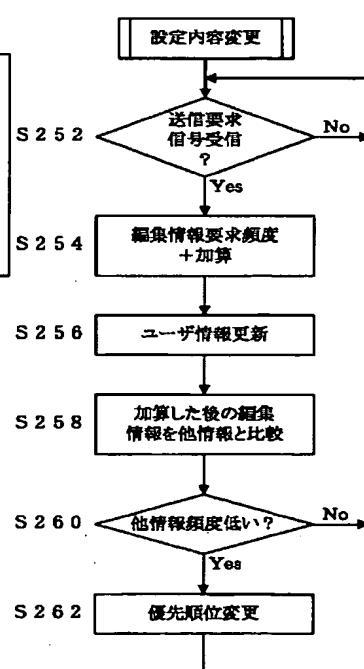
クリップの開始点・終了点を示すデータ構造

時 (0~99)	分 (0~59)	秒 (0~59)	フレーム (0~29)
1バイト	1バイト	1バイト	1バイト

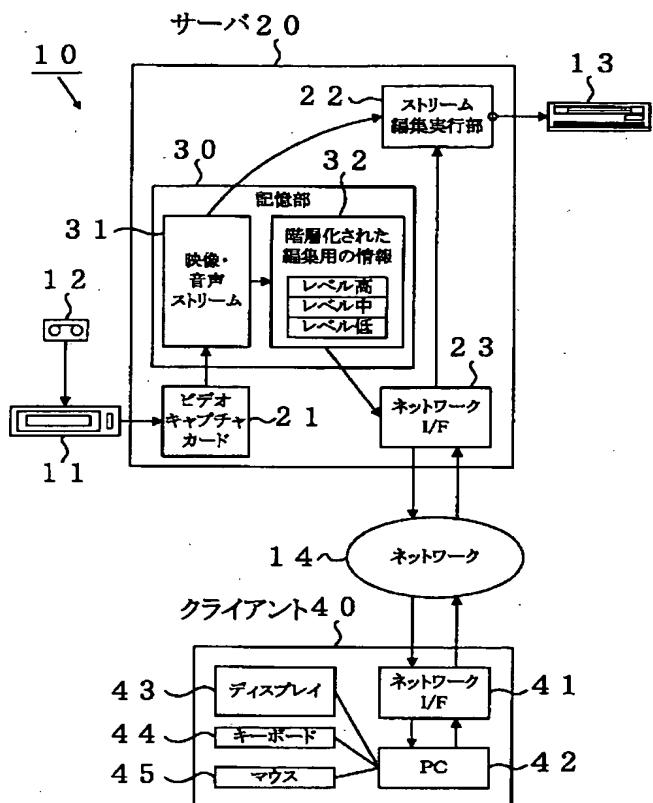
【図5】



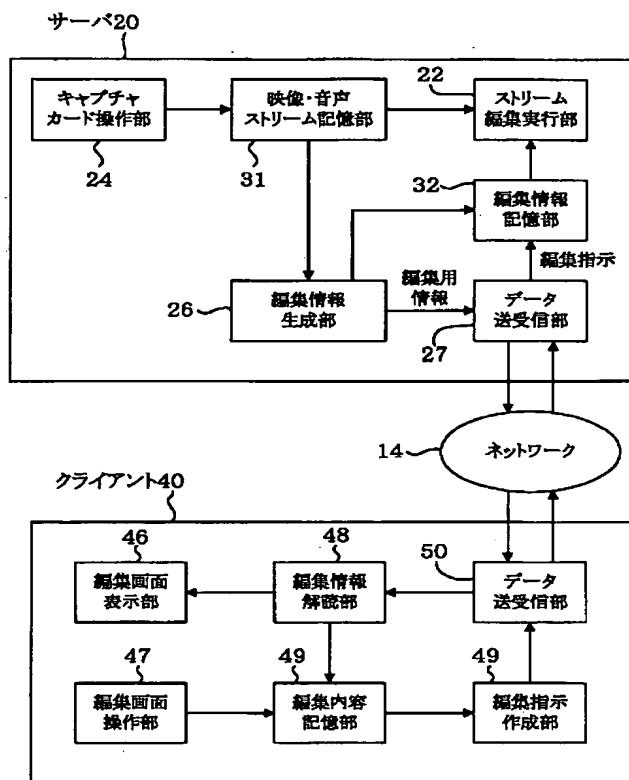
【図13】



( 1)

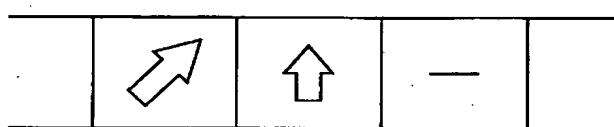


〔图2〕



(图 8)

(A)



(B)

## 動きベクトル座標

45

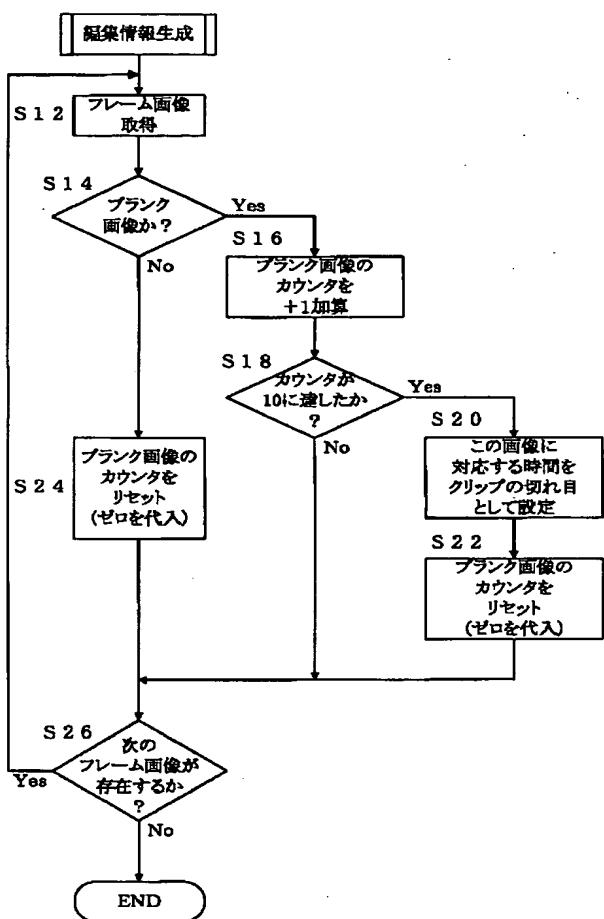
(0,1)	(10,10)	(3,6)	
(0,1)	なし	(-8,-3)	
(-1,0)	なし	(-8,0)	...
			⋮

【図3】

3.2

3 2 a	サムネイル画像
3 2 b	2値画像
3 2 c	画像の一部
3 2 d	線画
3 2 e	クローズド・キャプション
3 2 f	テープラベル情報
3 2 g	音情報
3 2 h	クリップ開始点・終了点
3 2 i	クリップ数・番号
3 2 j	フレーム数
3 2 k	シーンチェンジ点数
3 2 m	動きベクトル
3 2 n	単色画像

【図7】



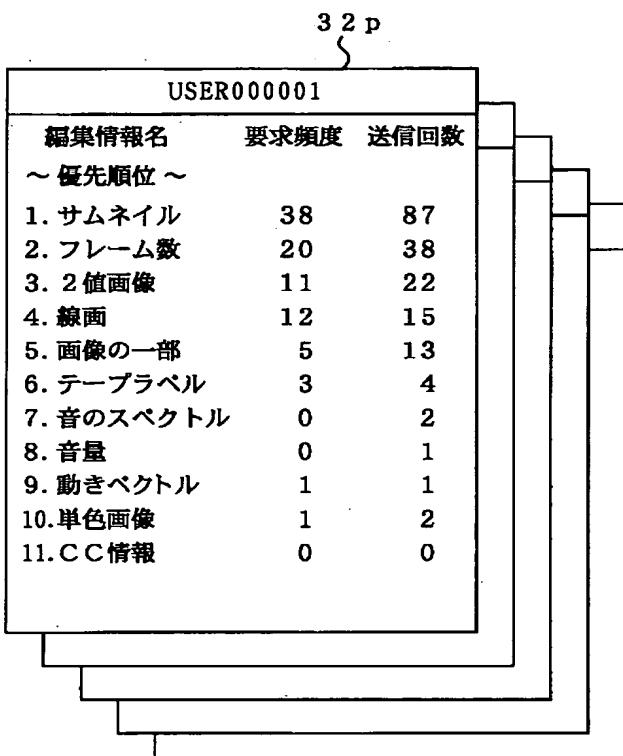
【図10】

4.3

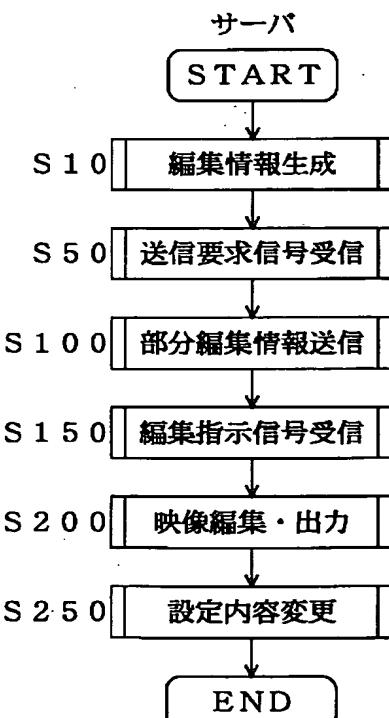
学習状況（上位10情報）									
1. フレーム数	6. 単色								
2. サムネイル	7. 二値画像								
3. テーブラベル	8. 線画								
4. 画像の一部	9. G.O.P構造								
5. 音量	10. 動きベクトル								

上記学習を  クリヤー  保存 する =>

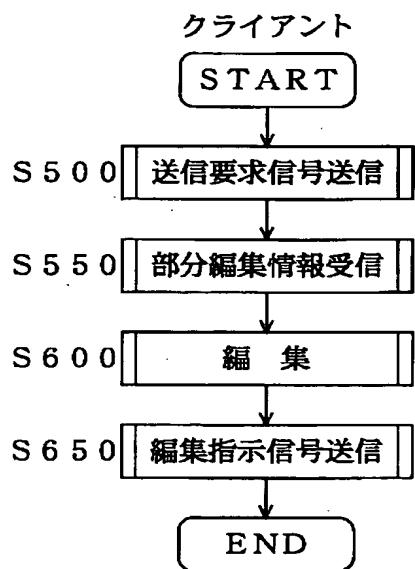
【図9】



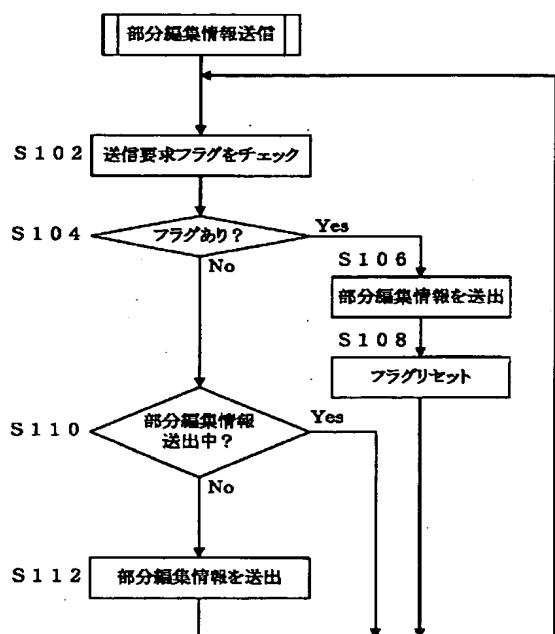
【図11】



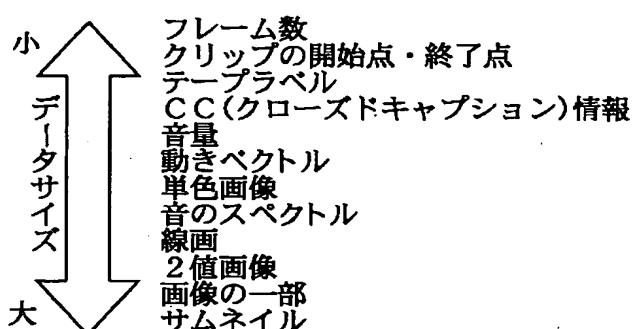
【図12】



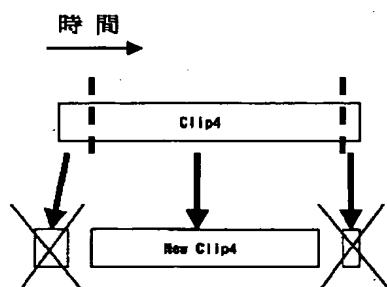
【図14】



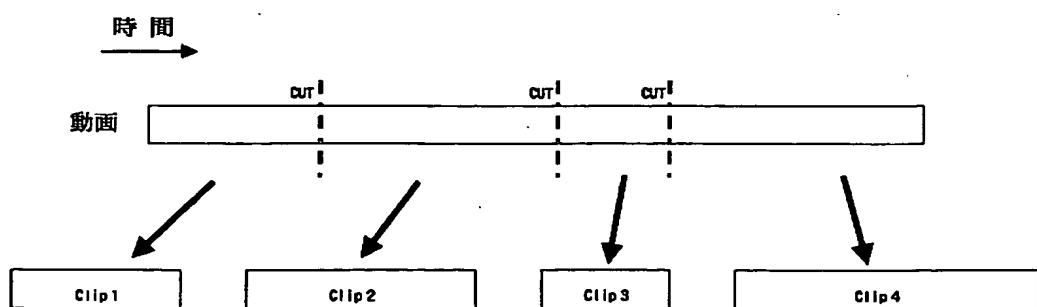
【図15】



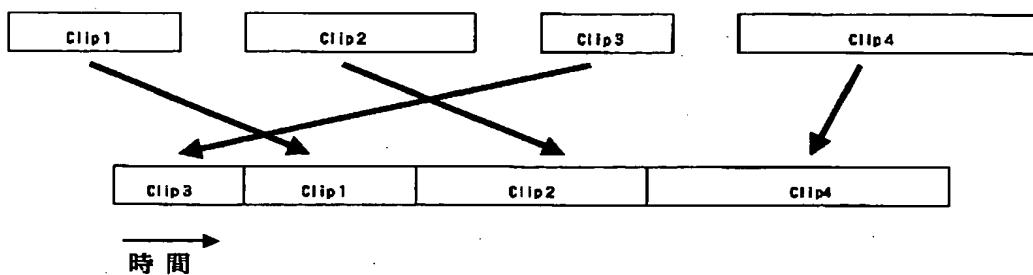
【図18】



【図16】



【図17】



## 【手続補正書】

【提出日】平成12年10月20日(2000.10.20)

## 【手続補正1】

【補正対象書類名】明細書

【補正対象項目名】0024

【補正方法】変更

## 【補正内容】

【0024】つまり、サーバは、該当する映像をクライアントへ送信するのではなく、自身に接続された映像記\*

\* 憶装置へ送信するため、クライアントは、サーバから送信された映像を記憶する必要がない。したがって、クライアントは、サーバから送信された映像を受信するための通信コストがかからないとともに、受信した映像を記憶するための記憶領域を確保しておく必要もない。また、クライアントは、上記映像記憶装置に備えられた記憶媒体（たとえば、DVD-R、DVD-RW、DVD-RAM、ビデオテープなど）をサーバから受け取ることにより、自身が編集した映像入手できる。

## フロントページの続き

(72)発明者 石本 関  
名古屋市瑞穂区苗代町15番1号 ブラザー  
工業株式会社内

(72)発明者 田川 典生  
名古屋市瑞穂区苗代町15番1号 ブラザー  
工業株式会社内

(72)発明者 市川 恭久  
名古屋市瑞穂区苗代町15番1号 ブラザー  
工業株式会社内  
F ターム(参考) 5C053 FA14 FA24 FA30 JA30 KA01  
KA24 LA11 LA14  
5C064 BA07 BC10 BC16 BC20 BC25  
BD01 BD08  
5D110 CA05 CA06 CD02 CD05 CD15